

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

United States Patent and Trademark
Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 18 June 1998 (18.06.98)	
International application No. PCT/EP97/05441	Applicant's or agent's file reference PCT 800-031/aw
International filing date (day/month/year) 02 October 1997 (02.10.97)	Priority date (day/month/year) 02 October 1996 (02.10.96)
Applicant BUJARD, Hermann et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:

23 April 1998 (23.04.98)

☐ in a notice effecting later election filed with the International Bureau on:
2. The election ☒ was
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

A. Addae-Ruesch

Telephone No.: (41-22) 338.83.38

THIS PAGE BLANK (USPTO)

09-2698721

PCT/EP97/05441

PCT COOPERATION TREATY

PCT

NOTIFICATION CONCERNING
DOCUMENT TRANSMITTED

From the INTERNATIONAL BUREAU

To:

United States Patent and Trademark
Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year)

16 April 1999 (16.04.99)

International application No.

PCT/EP97/05441

International filing date (day/month/year)

02 October 1997 (02.10.97)

Applicant

BUJARD, Hermann et al

The International Bureau transmits herewith the following documents and number thereof:

_____ copy of the English translation of the international preliminary examination report (Article 36(3)(a))

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

R. E. Stoffel

Telephone No.: (41-22) 338.83.38

THIS PAGE BLANK (USPTO)

VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS

PCT

INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

(Artikel 36 und Regel 70 PCT)

REC'D 02 FEB 1999

WIPO PCT



Aktenzeichen des Anmelders oder Anwalts PCT 800-031/ps	WEITERES VORGEHEN siehe Mitteilung über die Übersendung des internationalen vorläufigen Prüfungsberichts (Formblatt PCT/IPEA/416)	
Internationales Aktenzeichen PCT/EP97/05441	Internationales Anmeldedatum (Tag/Monat/Jahr) 02/10/1997	Priority date (Tag/Monat/Jahr) 02/10/1996
Internationale Patentklassifikation (IPK) oder nationale Klassifikation und IPK C12N15/30		
Anmelder BUJARD, Herman et al.		

- Dieser internationale vorläufige Prüfungsbericht wurde von der mit der internationalen vorläufigen Prüfung beauftragten Behörde erstellt und wird dem Anmelder gemäß Artikel 36 übermittelt.
- Dieser BERICHT umfaßt insgesamt 8 Blätter einschließlich dieses Deckblatts.
 - ☒ Außerdem liegen dem Bericht ANLAGEN bei; dabei handelt es sich um Blätter mit Beschreibungen, Ansprüchen und/oder Zeichnungen, die geändert wurden und diesem Bericht zugrunde liegen, und/oder Blätter mit vor dieser Behörde vorgenommenen Berichtigungen (siehe Regel 70.16 und Abschnitt 607 der Verwaltungsrichtlinien zum PCT).

Diese Anlagen umfassen insgesamt 6 Blätter.

3. Dieser Bericht enthält Angaben zu folgenden Punkten:

- I ☒ Grundlage des Berichts
- II ☐ Priorität
- III ☐ Keine Erstellung eines Gutachtens über Neuheit, erfinderische Tätigkeit und gewerbliche Anwendbarkeit
- IV ☒ Mangelnde Einheitlichkeit der Erfindung
- V ☒ Begründete Feststellung nach Artikel 35(2) hinsichtlich der Neuheit, der erfinderischen Tätigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erklärungen zur Stützung dieser Feststellung
- VI ☐ Bestimmte angeführte Unterlagen
- VII ☐ Bestimmte Mängel der internationalen Anmeldung
- VIII ☐ Bestimmte Bemerkungen zur internationalen Anmeldung

Datum der Einreichung des Antrags 23/04/1998	Datum der Fertigstellung dieses Berichts 28.01.99
Name und Postanschrift der mit der internationalen vorläufigen Prüfung beauftragten Behörde  Europäisches Patentamt D-80298 München Tel. (+49-89) 2399-0, Tx: 523656 epmu d Fax: (+49-89) 2399-4465	Bevollmächtigter Bediensteter Kalsner, I Telefon (+49-89) 2399-8708 

THIS PAGE BLANK (USPTO)

INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

Internationales Aktenzeichen PCT/EP97/05441

I. Grundlage des Berichts

1. Dieser Bericht wurde erstellt auf der Grundlage (*Ersatzblätter, die dem Anmeldeamt auf eine Aufforderung nach Artikel 14 hin vorgelegt wurden, gelten im Rahmen dieses Berichts als "ursprünglich eingereicht" und sind ihm nicht beigelegt, weil sie keine Änderungen enthalten.*):

Beschreibung, Seiten:

1-24 ursprüngliche Fassung

Patentansprüche, Nr.:

1-41 eingegangen am 10/12/1998 mit Schreiben vom 10/12/1998

Zeichnungen, Blätter:

1/16-16/16 ursprüngliche Fassung

2. Aufgrund der Änderungen sind folgende Unterlagen fortgefallen:

- ☐ Beschreibung, Seiten:
- ☐ Ansprüche, Nr.:
- ☐ Zeichnungen, Blatt:

3. ☐ Dieser Bericht ist ohne Berücksichtigung (von einigen) der Änderungen erstellt worden, da diese aus den angegebenen Gründen nach Auffassung der Behörde über den Offenbarungsgehalt in der ursprünglich eingereichten Fassung hinausgehen (Regel 70.2(c)):

4. Etwaige zusätzliche Bemerkungen:

IV. Mangelnde Einheitlichkeit der Erfindung

1. Auf die Aufforderung zur Einschränkung der Ansprüche oder zur Zahlung zusätzlicher Gebühren hat der Anmelder:

- ☐ die Ansprüche eingeschränkt.
- ☒ zusätzliche Gebühren entrichtet.
- ☐ zusätzliche Gebühren unter Widerspruch entrichtet.
- ☐ weder die Ansprüche eingeschränkt noch zusätzliche Gebühren entrichtet.

THIS PAGE BLANK (USPTO)

INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

Internationales Aktenzeichen PCT/EP97/05441

2. ☐ Die Behörde hat festgestellt, daß das Erfordernis der Einheitlichkeit der Erfindung nicht erfüllt ist, und hat gemäß Regel 68.1 beschlossen, den Anmelder nicht zur Einschränkung der Ansprüche oder zur Zahlung zusätzlicher Gebühren aufzufordern.
3. Die Behörde ist der Auffassung, daß das Erfordernis der Einheitlichkeit der Erfindung nach den Regeln 13.1, 13.2 und 13.3
- ☐ erfüllt ist
- ☒ aus folgenden Gründen nicht erfüllt ist:
siehe Beiblatt
4. Daher wurde zur Erstellung dieses Berichts eine internationale vorläufige Prüfung für folgende Teile der internationalen Anmeldung durchgeführt:
- ☒ alle Teile.
- ☐ die Teile, die sich auf die Ansprüche Nr. beziehen.

V. Begründete Feststellung nach Artikel 35(2) hinsichtlich der Neuheit, der erfinderischen Tätigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erklärungen zur Stützung dieser Feststellung

1. Feststellung

Neuheit (N)	Ja: Ansprüche	1-25, 27, 28, 30-36, teilweise 38, 39
	Nein: Ansprüche	26, 29, 37, 40, 41, teilweise 38, 39
Erfinderische Tätigkeit (ET)	Ja: Ansprüche	1-25, 33-36, teilweise 38, 39
	Nein: Ansprüche	26-32, 37, 40, 41, teilweise 38, 39
Gewerbliche Anwendbarkeit (GA)	Ja: Ansprüche	1-41
	Nein: Ansprüche	

2. Unterlagen und Erklärungen

siehe Beiblatt

THIS PAGE BLANK (USPTO)

Zu Abschnitt IV: Mangelnde Einheitlichkeit der Erfindung

Die Internationale Prüfungsbehörde (IPEA) schließt sich dem von der Internationalen Recherchenbehörde (ISA) wegen mangelnder Einheitlichkeit vorgebrachten Einwand an.

Eine internationale Anmeldung darf sich nur auf eine einzige Erfindung, oder auf eine Gruppe von Erfindungen beziehen, die so zusammenhängen, daß sie eine einzige allgemeine erfinderische Idee verwirklichen.

Das Erfordernis der Einheitlichkeit der Erfindung nach Art. 34(3)a) und Regel 13.1 PCT ist nur erfüllt, wenn zwischen den Erfindungen ein technischer Zusammenhang besteht, der in einem oder mehreren gleichen oder entsprechenden besonderen technischen Merkmalen zum Ausdruck kommt. Unter dem Begriff "besondere technische Merkmale" sind diejenigen technischen Merkmale zu verstehen, die einen Beitrag jeder beanspruchten Erfindung als Ganzes zum Stand der Technik bestimmen.

Die Argumente betreffend den Mangel von Einheitlichkeit, die bereits von der ISA vorgebracht wurden, werden *mutatis mutandis* von der IPEA übernommen. Wie schon von der ISA ausgeführt, wurden die folgenden zwei Erfindungen identifiziert:

Erfindung I: Ansprüche 1-35, 37-41 (teilweise):

DNA-Sequenz, die das gp190/MSP-1-Oberflächenprotein von *Plasmodium* kodiert, Verfahren zur Herstellung des Oberflächenproteins, Wirtsorganismen, die diese Sequenz enthalten, Verwendung des gp190/MSP-1-Oberflächenproteins zur Immunisierung, Vektor besagte DNA-Sequenz enthaltend, Impfstoff das gp190/MSP-1-Oberflächenprotein oder die dafür kodierende DNA enthaltend.

Erfindung II: Ansprüche 36, 37-41 (teilweise):

THIS PAGE BLANK (USPTO)

Verfahren zur Stabilisierung von Gensequenzen durch
Verringerung des AT-Gehaltes der Sequenz, stabilisiertes Gen,
Vektor dieses Gen enthaltend, Impfstoff diesen Vektor enthaltend.

Zwischen den beiden obengenannten Erfindungen ist kein gemeinsames
"besonderes technisches Merkmal" im Sinne von Regel 13.1 und 13.2 PCT zu
erkennen. Das Erfordernis der Einheitlichkeit ist somit nicht erfüllt.

Da der Anmelder der Aufforderung zur Bezahlung einer zweiten Prüfungsgebühr
nachkam, wurden beide Erfindungen geprüft.

**Zu Abschnitt V: Begründete Feststellung hinsichtlich Neuheit, erfinderischer
Tätigkeit und gewerblicher Anwendbarkeit**

I) Erfindung I: Ansprüche 1-35 und teilweise 37-41

I.1) Dokumente

D1...WO-A-9428930

D2...EP-A-0154454

I.2) Neuheit und erfinderische Tätigkeit

- a. D1 beschreibt ein Herstellungsverfahren für das vollständige gp190/MSP1-Protein von *Plasmodium falciparum*. Die Herstellung des Proteins erfolgte ausgehend von vier Plasmiden mit Teilsequenzen der für das Protein kodierenden Nukleinsäure, die zusammen für die vollständige Proteinsequenz kodieren. Die Teilstücke wurden fusioniert und in *Vaccinia* exprimiert. Die Expression von gp190 nach Infektion von HeLa-Zellen wurde mittels Immunopräzipitations-Analyse bestätigt. Desweiteren werden Impfstoffe, die das gp190-Oberflächenprotein enthalten beansprucht.

D2 offenbart die vollständige Nukleinsäure- und Proteinsequenz von gp190 aus *Plasmodium falciparum* (Abb. 1) und eine Zelle, die diese Nucleinsäuresequenz in einem Vektor enthält (Anspruch 14). Weiters werden Impfstoffe beansprucht, die

THIS PAGE BLANK (USPTO)

das gp190-Protein oder das gp190 und mindestens einen anderen anti-Malaria-Impfstoff enthalten.

- b. **Ansprüche 1 und 17**, sowie die davon direkt oder indirekt abhängigen **Ansprüche 2-16, 18-25, 33-35 und teilweise 38 und 39** erfüllen die Erfordernisse von Art. 33(2)(3) PCT.

Anspruch 1 bezieht sich auf ein Herstellungsverfahren für das vollständige gp190/MSP1-Protein von *Plasmodium falciparum*, das dadurch gekennzeichnet ist, daß das vollständige Gen für gp190/MSP1 in einem geeigneten System exprimiert wird, und daß der AT-Gehalt der zugrunde liegenden, exprimierten DNA-Sequenz gegenüber der natürlich vorkommenden Sequenz erniedrigt ist.

D1 beschreibt zwar die Klonierung und Expression von gp190/MSP1, jedoch ist dem Dokument kein Hinweis zu entnehmen, daß die DNA-Sequenz bezüglich ihres AT-Gehaltes modifiziert werden sollte. Wie aus der Beschreibung der vorliegenden Anmeldung hervorgeht, verleiht die Reduzierung des AT-Gehaltes der DNA-Sequenz eine höhere Stabilität. Diese Tatsache ist weder in D1 noch in den anderen im internationalen Recherchebericht zitierten Dokumenten offenbart oder daraus ableitbar.

Insofern als **Ansprüche 2-25, 33-35 und teilweise 38 und 39** direkt oder indirekt von Anspruch 1 abhängig sind erfüllen auch sie die Kriterien von Art. 33(2)(3) PCT.

- c. **Anspruch 26**, der sich auf einen Wirtsorganismus bezieht, der die vollständige Nukleinsäuresequenz gemäß einem der Ansprüche 17-25 für das gp190/MSP1-Oberflächenprotein und/oder das vollständige Protein enthält, und **Anspruch 29**, der den Wirtsorganismus genauer definiert (HeLa-Zellen), entsprechen angesichts der Offenbarungen von D1 (Herstellung des vollständigen gp190-Oberflächenprotein in Wirtsorganismen) nicht den Erfordernissen von Art. 33(2) PCT.

Darüber hinaus steht der natürlich vorkommende Wirtsorganismus *Plasmodium falciparum*, der ja zweifellos das vollständige gp190-Oberflächenprotein enthält,

THIS PAGE BLANK (USPTO)

Anspruch 26 neuheitsschädlich entgegen.

- d. **Ansprüche 27, 28 und 30-32**, die den Wirtsorganismus von Anspruch 26 näher definieren, erfüllen zwar die Erfordernisse von Art. 33(2) PCT, nicht jedoch die Erfordernisse von Art. 33(3) PCT, da die Bereitstellung verschiedener bekannter Wirtsorganismen zur Expression eines bekannten und klonierten Proteins für den Fachmann offensichtlich ist.
- e. **Ansprüche 40 und 41** entsprechen den Kriterien von Art. 33(2) PCT nicht, da ein bekannter Impfstoff nicht dadurch neu wird, daß er auf eine andere Weise hergestellt wird.

II) Erfindung II: Ansprüche 36, 37 und teilweise 38-41

II.1) Dokumente

D4...EP-A-0385962

D5...EP-A-0359472

II.2) Neuheit und erfinderische Tätigkeit

- a. D4 und D5 beschreiben synthetische DNA-Sequenzen, die für *Bacillus thuringensis* Toxin kodieren, wobei der AT-Gehalt der Sequenzen verringert wurde, sowie die Herstellung dieser Nukleinsäuren (D4, Beispiele 2 und 3, Abb. 2-4; D5, S. 21, Zeilen 34-44, Tab. 3). Die synthetischen Oligonukleotide wurden in verschiedenen Pflanzen exprimiert.
- b. **Anspruch 36** erfüllt die Kriterien von Art. 33(2)(3) PCT, da ein Verfahren zur Stabilisierung von Gen-Sequenzen, das darauf beruht, den AT-Gehalt einer Sequenz zu verringern und damit dem Gen eine höhere Stabilität zu verleihen, als solches nicht im vorhandenen Stand der Technik beschrieben oder nahegelegt ist.
- c. **Ansprüche 37 und teilweise 38 und 39** können gegenüber D4 und D5 nicht als neu im Sinne von Art. 33(2) PCT anerkannt werden, da man annehmen kann, daß die in den beiden Dokumenten beschriebenen DNA-Sequenzen, deren AT-Gehalt

THIS PAGE BLANK (USPTO)

—
reduziert wurde, stabiler sind als unmodifizierten Ausgangsprodukte, auch wenn eine erhöhte Stabilität nicht der ausdrückliche Beweggrund für die Modifikationen war.

- d. **Ansprüche 40 und 41**, insofern sie sich auf Anspruch 37 beziehen, entsprechen nicht den Erfordernissen von Art. 33(3) PCT, da die Bereitstellung von Impfstoffen, die nicht als neu anerkannte Vektoren enthalten, als offensichtlich anzusehen ist.

THIS PAGE BLANK (USPTO)

PCT/EP97/05441
Anmelder. BUJARD, Hermann u.a.
PCT 800-01996/co
10.12.1998

(Neue) Patentansprüche

1. Herstellungsverfahren für das vollständige gp190/MSP1-Protein von Plasmodium, insbesondere *Plasmodium falciparum*, **dadurch gekennzeichnet**, daß das vollständige Gen für gp190/MSP1 in einem geeigneten System, vorzugsweise einem Wirtsorganismus, exprimiert wird, und daß der AT-Gehalt der dem Protein zugrunde liegenden, exprimierten DNA-Sequenz gegenüber der natürlich vorkommenden Sequenz erniedrigt ist.
2. Rekombinantes Herstellungsverfahren gemäß Anspruch 1, **dadurch gekennzeichnet**, daß der Synthese der dem Protein zugrunde liegenden DNA-Sequenz die DNA-Sequenz des *P. falciparum*-Stammes FCB-1 zugrunde gelegt wird.
3. Rekombinantes Herstellungsverfahren gemäß Anspruch 2, **dadurch gekennzeichnet**, daß der AT-Gehalt von 74% auf 55% erniedrigt ist.
4. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz einschließlich Signalpeptid und Ankersignal kodiert.
5. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals kodiert.
6. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende

GEÄNDERTES BLATT
GEÄNDERTES BLATT
GEÄNDERTES BLATT

THIS PAGE BLANK (USPTO)

Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals und des Signalpeptides kodiert.

7. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 6, **dadurch gekennzeichnet**, daß es folgende Schritte umfaßt:
 - (a) Entwurf der zu synthetisierenden DNA-Sequenz aus *P. falciparum* FCB-1, wobei eine DNA-Sequenz mit den im menschlichen Genom üblichen Codon-Häufigkeiten unter Erhalt der Aminosäuresequenz des FCB-1 Proteins hergestellt werden sollte,
 - (b) Einteilung der entworfenen Sequenz in überlappende Regionen, vorzugsweise in Regionen p83, p31, p36, gp30 und gp19,
 - (c) Synthese von Desoxyoligonukleotiden, die jeweils mindestens die gesamte Länge einer Region abdecken,
 - (d) Synthese der kodierenden Regionen für gp19, gp30, p36 und p31 durch PCR und Synthese der kodierenden Region für p83 durch Fusion aus zwei etwa 1200bp umfassenden Sequenzen,
 - (e) einzelne Klonierung der kodierenden Sequenzen
 - (f) Fusion des gesamten Gens und
 - (g) Expression in einem geeigneten Expressionssystem.
8. Rekombinantes Herstellungsverfahren nach Anspruch 7, **dadurch gekennzeichnet**, daß die in Schritt (c) synthetisierten Desoxyoligonukleotide durchschnittlich 120 Nukleotide lang sind und die benachbarten Sequenzen jeweils um ca 20 Basen überlappen.
9. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor dPS56, RBSII verwendet wird.

THIS PAGE BLANK (USPTO)

10. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor pBi-5 verwendet wird.
11. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor ppTMCS verwendet wird.
12. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in *E. coli* exprimiert wird.
13. Rekombinantes Herstellungsverfahren nach Anspruch 12, **dadurch gekennzeichnet**, daß der verwendete *E. coli*-Stamm der Repressor-produzierende Stamm *E. coli* DH5alphaZ1 ist.
14. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in HeLa-Zellen exprimiert wird.
15. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in CHO-Zellen exprimiert wird.
16. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in *Toxoplasma gondii* oder *Leishmania* exprimiert wird.
17. Vollständige, zur Expression geeignete DNA-Sequenz des gp190/MSP1-Oberflächenproteins von Plasmodium, insbesondere *P. falciparum*, erhältlich durch das re-

THIS PAGE BLANK (USPTO)

kombinierte Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 16. —

18. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 17, **dadurch gekennzeichnet**, daß sie nicht für das Ankersignal kodiert.
19. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 18, **dadurch gekennzeichnet**, daß sie weder für das Ankersignal noch für das Signalpeptid kodiert.
20. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 19, **dadurch gekennzeichnet**, daß sie eine Sequenz für am N-Terminus vorliegende 11 zusätzliche Aminosäuren, davon 6 Histidine, umfaßt.
21. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 20, **dadurch gekennzeichnet**, daß die Sequenz keine erkennbaren "splice donor"- und "splice acceptor"-Signale enthält.
22. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 21, **dadurch gekennzeichnet**, daß die Sequenz keine größeren GC-reichen Sequenzen enthält.
23. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 22, **dadurch gekennzeichnet**, daß die Sequenz keine Erkennungssignale für Restriktionsenzyme, welche Sequenzen von sechs oder mehr Basenpaaren erkennen, enthält.
24. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 23, **dadurch gekennzeichnet**, daß die Sequenz für Erkennungssignale bestimmter Restriktionsnukleasen in Regionen, die die nach der Prozessierung des Pro-

THIS PAGE BLANK (USPTO)

teins entstehenden Domänen trennen, einmal vorkommende Schnittstellen für Restriktionsendonukleasen enthält.

25. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 24, **dadurch gekennzeichnet**, daß die Sequenz an ihren beiden Enden Schnittstellen für Restriktionsendonukleasen aufweist, die in der übrigen Sequenz und in einem zu verwendenden Vektor nicht vorkommen.
26. Wirtsorganismus, der die vollständige Nukleinsäuresequenz gemäß einem der Ansprüche 17 bis 25 für das gp190/MSP1-Oberflächenprotein und/oder das vollständige Protein enthält.
27. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus *E. coli* ist.
28. Wirtsorganismus gemäß Anspruch 27, **dadurch gekennzeichnet**, daß der *E. coli*-Stamm der Repressor-produzierende *E. coli*-Stamm DH5alphaZ1 ist.
29. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus HeLa-Zellen sind.
30. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus CHO-Zellen sind.
31. Wirtsorganismus gemäß Anspruch 29 oder 30, **dadurch gekennzeichnet**, daß die Wirtszellen konstitutiv tTA synthetisieren.
32. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus *Toxoplasma gondii*, *Leishmania*, Baculoviren, Adenoviren oder Hefen vorgesehen sind.

THIS PAGE BLANK (USPTO)

33. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur aktiven Immunisierung gegen Malaria.
34. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur Herstellung von monoklonalen, zur passiven Immunisierung geeigneten Antikörpern.
35. Verwendung einer gemäß der Ansprüche 1 bis 16 hergestellten DNA-Sequenz zur Herstellung einer Vakzine auf Nukleinsäurebasis.
36. Verfahren zur Stabilisierung von Gen-Sequenzen, **dadurch gekennzeichnet**, daß der AT-Gehalt der Sequenz verringert wird.
37. Stabilisiertes Gen, **dadurch gekennzeichnet**, daß es einen geringeren AT-Gehalt aufweist als das nicht stabilisierte Gen.
38. Vektor enthaltend eine DNA-Sequenz nach einem der Ansprüche 17 bis 25 und/oder 37.
39. Wirtszelle enthaltend einen Vektor nach Anspruch 38.
40. Impfstoff enthaltend ein Protein, hergestellt nach einem Verfahren nach einem der Ansprüche 1-16 und/oder eine DNA-Sequenz nach einem der Ansprüche 17-25 und/oder einen Wirt nach einem der Ansprüche 26-31 und/oder einen Vektor nach Anspruch 38.
41. Impfstoff nach Anspruch 40, **dadurch gekennzeichnet**, daß er weitere Immunität hervorrufende Produkte aus Plasmodium, insbesondere *P. falciparum*, enthält.

THIS PAGE BLANK (USPTO)

Translation

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference PCT 800-031/ps	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/EP97/05441	International filing date (day/month/year) 02 October 1997 (02.10.1997)	Priority date (day/month/year) 02 October 1996 (02.10.1996)
International Patent Classification (IPC) or national classification and IPC C12N 15/30, C07K 14/445, C12N 15/62, A61K 39/015, 31/70, C07H 21/00, C12N 1/21, 5/10, 1/11, 15/67		
Applicant BUJARD, Hermann		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of <u>8</u> sheets, including this cover sheet. <input checked="" type="checkbox"/> This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT). These annexes consist of a total of <u>6</u> sheets.
3. This report contains indications relating to the following items: I <input checked="" type="checkbox"/> Basis of the report II <input type="checkbox"/> Priority III <input type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicability IV <input checked="" type="checkbox"/> Lack of unity of invention V <input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement VI <input type="checkbox"/> Certain documents cited VII <input type="checkbox"/> Certain defects in the international application VIII <input type="checkbox"/> Certain observations on the international application

Date of submission of the demand 23 April 1998 (23.04.1998)	Date of completion of this report 28 January 1999 (28.01.1999)
Name and mailing address of the IPEA/EP European Patent Office D-80298 Munich, Germany Facsimile No. 49-89-2399-4465	Authorized officer Telephone No. 49-89-2399-0

THIS PAGE BLANK (USPTO)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/EP97/05441

I. Basis of the report

1. This report has been drawn on the basis of *(Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.)*:

- ☐ the international application as originally filed.
- ☒ the description, pages 1 - 24, as originally filed,
 pages _____, filed with the demand,
 pages _____, filed with the letter of _____,
 pages _____, filed with the letter of _____.
- ☒ the claims, Nos. _____, as originally filed,
 Nos. _____, as amended under Article 19,
 Nos. _____, filed with the demand,
 Nos. 1 - 41, filed with the letter of 10 December 1998 (10.12.1998),
 Nos. _____, filed with the letter of _____.
- ☒ the drawings, sheets/fig 1/16-16/16, as originally filed,
 sheets/fig _____, filed with the demand,
 sheets/fig _____, filed with the letter of _____,
 sheets/fig _____, filed with the letter of _____.

2. The amendments have resulted in the cancellation of:

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheets/fig _____

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).

4. Additional observations, if necessary:

THIS PAGE BLANK (USPTO)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/EP97/05441

IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

- ☐ restricted the claims.
- ☒ paid additional fees.
- ☐ paid additional fees under protest.
- ☐ neither restricted nor paid additional fees.

2. ☐ This Authority found that the requirement of unity of invention is not complied with and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

- ☐ complied with.
- ☒ not complied with for the following reasons:

See Supplemental Box

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

- ☒ all parts.
- ☐ the parts relating to claims Nos. _____

THIS PAGE BLANK (USPTO)

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: IV.

The International Preliminary Examining Authority (IPEA) agrees with the objection regarding lack of unity raised by the International Searching Authority (ISA).

An international application can only relate to a single invention or to a group of inventions that are related so as to embody a single general inventive idea.

The requirement of unity of invention according to PCT Article 34(3)(a) and Rule 13.1 is only satisfied if a technical connection exists between the inventions that is expressed in one or more of the same or corresponding special technical features. The term "special technical features" means those technical features that define a contribution which each claimed inventions, considered as a whole, makes over the prior art.

The arguments concerning lack of unity that already have been raised by the ISA are applied *mutatis mutandis* by the IPEA. As already stated by the ISA, the following two inventions have been identified:

Invention I: Claims 1-35, 37-41 (in part):

a DNA sequence, which codes the gp190/MSP-1 cell-surface protein of *Plasmodium*, a process for producing the cell-surface protein, host

THIS PAGE BLANK (USPTO)

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: IV.

organisms containing the sequence, use of the gp190/MSP-1 cell-surface protein for immunization, vectors containing said DNA sequence, vaccine containing gp190/MSP-1 cell-surface protein or the DNA coding for same.

Invention II: Claims 36, 37-41 (in part):

a process for stabilizing gene sequences through lowering the AT content of the sequence, a stabilized gene, a vector containing this gene, a vaccine containing this vector.

There is no common "special technical feature" in the sense of PCT Rule 13.1 and 13.2 between the two inventions cited above. Hence, the requirement of unity is not satisfied.

Because the applicant agreed to payment of a second examining fee, both inventions have been examined.

THIS PAGE BLANK (USPTO)

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	<u>1-25, 27, 28, 30-36, in part 38, 39</u>	YES
	Claims	<u>26, 29, 37, 40, 41, in part 38, 39</u>	NO
Inventive step (IS)	Claims	<u>1-25, 33-36, in part 38, 39</u>	YES
	Claims	<u>26-32, 37, 40, 41, in part 38, 39</u>	NO
Industrial applicability (IA)	Claims	<u>1-41</u>	YES
	Claims		NO

2. Citations and explanations

I) Invention 1: Claims 1-35 and in part 37-41

I.1) Documents

D1 = WO-A-9428930

D2 = EP-A-0154454

I.2) Novelty and Inventive Step

a. D1 describes a production process for the complete gp190/MSP1 protein of *Plasmodium falciparum*. The production of the protein is done starting with four plasmids with partial sequences of the nucleic acid coding for the protein, which together code for the entire protein sequence. The partial segments were fused and expressed in vaccinia. The expression of gp190 after infection by HeLa cells was confirmed by means of immunoprecipitation analyses. Moreover, vaccines are claimed which contain the gp190 cell-surface protein.

D2 discloses the complete sequence of nucleic acid and protein of gp190 from *Plasmodium falciparum*

THIS PAGE BLANK (USPTO)

(Fig. 1) and a cell that contains this nucleic acid sequence in a vector (Claim 14). In addition, vaccines are claimed that contain the gp190 protein or gp190 and at least one other anti-malarial vaccine.

- b. **Claims 1 and 17, as well as Claims 2-16, 18-25, 33-35 and in part 38 and 39, which directly or indirectly are dependent on the former, satisfy the requirements of PCT Article 33(2) and (3).**

Claim 1 relates to a production process for the complete gp190/MSP1 protein of *Plasmodium falciparum*, which is characterized in that a complete gene for gp190/MSP1 is expressed in a suitable system and that the AT content of the underlying expressed DNA sequence is lowered relative to the naturally occurring sequence.

D1 describes the cloning and expression of gp190/MSP1, but, there is no clear indication in the document that the DNA sequence should be modified with respect to its AT content. As the application makes clear, reduction of the AT content gives the DNA sequence greater stability. Neither D1 nor the documents cited in the international search report imply or disclose this fact.

Insofar as **Claims 2-25, 33-35 and in part 38 and 39** directly or indirectly depend on Claim 1, they too satisfy the criteria in PCT Article 33(2) and (3).

- c. **Claim 26, which relates to a host organism containing the entire nucleic acid sequence according to one of Claims 17-25 for the gp190/MSP1**

THIS PAGE BLANK (USPTO)

cell-surface protein and/or the entire protein, and Claim 29, which more closely defines the host organism (HeLa cells), do not satisfy the requirements of PCT Article 33(2) due to the disclosures by D1 (production of the complete gp190 cell-surface protein in host organisms).

Moreover, the naturally occurring host organism *Plasmodium falciparum*, which certainly contains the complete gp190 cell-surface protein, is prejudicial to the novelty of Claim 26.

- d. Claims 27, 28 and 30-32, which more closely define the host mechanism of Claim 26, in fact satisfy the requirements of PCT Article 33(2), not however the requirements of PCT Article 33(3) because the preparation of different known host organisms for expressing a known and cloned protein is obvious to a person skilled in the art.
- e. Claims 40 and 41 do not satisfy the criteria of PCT Article 33(2) because a known vaccine is not novel because it has been produced in another fashion.

II.1) Documents

D4 = EP-A-0385962

D5 = EP-A-0359472

II.2) Novelty and an Inventive Step

- a. D4 and D5 describe synthetic DNA sequences that code for *Bacillus thuringensis* toxin, where the AT content of the sequences has been reduced, and the production of these nucleic acids (D4, Examples 2

THIS PAGE BLANK (USPTO)

and 3, Figures 2-4; D5, page 21, lines 34-44, Table 3). The synthetic oligonucleotides were expressed in different plants.

- b. **Claim 36** satisfies the criteria of PCT Article 33(2) and (3) because a process for stabilizing gene sequences that depends upon reducing the AT content of a sequence and thereby giving the gene greater stability is as such neither disclosed nor suggested in the present prior art.
- c. **Claims 37 and in part 38 and 39** cannot be acknowledged as novel over D4 and D5 in the sense of PCT Article 33(2) because it can be assumed that the DNA sequences described in both documents, the AT content of which has been reduced, are more stable than unmodified starting products--even if the specific motivation for the modifications was not to increase stability.
- d. **Claims 40 and 41**, in so far as they refer to Claim 37, do not satisfy the requirements of PCT Article 33(3) because the preparation of vaccines that do not contain recognized vectors must be seen as obvious.

THIS PAGE BLANK (USPTO)

VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT
AUF DEM GEBIET DES PATENTWESENS

PCT

INTERNATIONALER RECHERCHENBERICHT

(Artikel 18 sowie Regeln 43 und 44 PCT)

Aktenzeichen des Anmelders oder Anwalts PCT 800-031/aw	WEITERES VORGEHEN siehe Mitteilung über die Übermittlung des internationalen Recherchenberichts (Formblatt PCT/ISA/220) sowie, soweit zutreffend, nachstehender Punkt 5	
Internationales Aktenzeichen PCT/EP 97/ 05441	Internationales Anmeldedatum (Tag/Monat/Jahr) 02/10/1997	(Frühestes) Prioritätsdatum (Tag/Monat/Jahr) 02/10/1996
Anmelder BUJARD, Hermann et al.		

Dieser internationale Recherchenbericht wurde von der Internationalen Recherchenbehörde erstellt und wird dem Anmelder gemäß Artikel 18 übermittelt. Eine Kopie wird dem Internationalen Büro übermittelt.

Dieser internationale Recherchenbericht umfaßt insgesamt 6 Blätter.

☒ Darüber hinaus liegt ihm jeweils eine Kopie der in diesem Bericht genannten Unterlagen zum Stand der Technik bei.

1. ☒ Bestimmte Ansprüche haben sich als nicht recherchierbar erwiesen (siehe Feld I).
2. ☒ Mangelnde Einheitlichkeit der Erfindung (siehe Feld II).
3. ☒ In der internationalen Anmeldung ist ein Protokoll einer Nucleotid- und/oder Aminosäuresequenz offenbart; die internationale Recherche wurde auf der Grundlage des Sequenzprotokolls durchgeführt,
 - ☐ das zusammen mit der internationalen Anmeldung eingereicht wurde.
 - ☒ das vom Anmelder getrennt von der internationalen Anmeldung vorgelegt wurde,
 - ☐ dem jedoch keine Erklärung beigefügt war, daß der Inhalt des Protokolls nicht über den Offenbarungsgehalt der internationalen Anmeldung in der eingereichten Fassung hinausgeht.
 - ☐ das von der Internationalen Recherchenbehörde in die ordnungsgemäße Form übertragen wurde.
4. Hinsichtlich der Bezeichnung der Erfindung
 - ☒ wird der vom Anmelder eingereichte Wortlaut genehmigt.
 - ☐ wurde der Wortlaut von der Behörde wie folgt festgesetzt.
5. Hinsichtlich der Zusammenfassung
 - ☒ wird der vom Anmelder eingereichte Wortlaut genehmigt.
 - ☐ wurde der Wortlaut nach Regel 38.2b) in der Feld III angegebenen Fassung von dieser Behörde festgesetzt. Der Anmelder kann der Internationalen Recherchenbehörde innerhalb eines Monats nach dem Datum der Absendung dieses internationalen Recherchenberichts eine Stellungnahme vorlegen.
6. Folgende Abbildung der Zeichnungen ist mit der Zusammenfassung zu veröffentlichen:
Abb. Nr. _____
 - ☐ wie vom Anmelder vorgeschlagen
 - ☐ weil der Anmelder selbst keine Abbildung vorgeschlagen hat.
 - ☐ weil diese Abbildung die Erfindung besser kennzeichnet.

☒ keine der Abb.

THIS PAGE BLANK (USPTO)

Feld I Bemerkungen zu den Ansprüchen, die sich als nicht recherchierbar erwiesen haben (Fortsetzung von Punkt 1 auf Blatt 1)

Gemäß Artikel 17(2)a) wurde aus folgenden Gründen für bestimmte Ansprüche kein Recherchenbericht erstellt:

1. ☒ Ansprüche Nr.
weil sie sich auf Gegenstände beziehen, zu deren Recherche die Behörde nicht verpflichtet ist, nämlich
Bemerkung : Obwohl die Ansprüche 33-35 sich auf ein Verfahren zur Behandlung des menschlichen/tierischen Körpers beziehen, wurde die Recherche durchgeführt und gründete sich auf die angeführten Wirkungen der Verbindung/Zusammensetzung.
2. ☐ Ansprüche Nr.
weil sie sich auf Teile der internationalen Anmeldung beziehen, die den vorgeschriebenen Anforderungen so wenig entsprechen, daß eine sinnvolle internationale Recherche nicht durchgeführt werden kann, nämlich
3. ☐ Ansprüche Nr.
weil es sich dabei um abhängige Ansprüche handelt, die nicht entsprechend Satz 2 und 3 der Regel 6.4 a) abgefaßt sind.

Feld II Bemerkungen bei mangelnder Einheitlichkeit der Erfindung (Fortsetzung von Punkt 2 auf Blatt 1)

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere Erfindungen enthält:

siehe Zusatzblatt

1. ☒ Da der Anmelder alle erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht auf alle recherchierbaren Ansprüche der internationalen Anmeldung.
2. ☐ Da für alle recherchierbaren Ansprüche die Recherche ohne einen Arbeitsaufwand durchgeführt werden konnte, der eine zusätzliche Recherchegebühr gerechtfertigt hätte, hat die Internationale Recherchenbehörde nicht zur Zahlung einer solchen Gebühr aufgefordert.
3. ☐ Da der Anmelder nur einige der erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht nur auf die Ansprüche der internationalen Anmeldung, für die Gebühren entrichtet worden sind, nämlich auf die Ansprüche Nr.
4. ☐ Der Anmelder hat die erforderlichen zusätzlichen Recherchegebühren nicht rechtzeitig entrichtet. Der internationale Recherchenbericht beschränkt sich daher auf die in den Ansprüchen zuerst erwähnte Erfindung; diese ist in folgenden Ansprüchen erfaßt:

Bemerkungen hinsichtlich eines Widerspruchs

- ☐ Die zusätzlichen Gebühren wurden vom Anmelder unter Widerspruch gezahlt.
- ☒ Die Zahlung zusätzlicher Gebühren erfolgte ohne Widerspruch.

THIS PAGE BLANK (USPTO)

WEITERE ANGABEN

PCT/ISA/ 210

1. Ansprüche: 1-35, teilweise 37-40

DNA-Sequenz das gp190/MSP-1 Oberflächenprotein von Plasmodium kodierend, Wirtsorganismus diese Sequenz enthaltend, Verwendung des gp190/MSP-1 Oberflächenproteins oder der dafür kodierenden DNA zur Immunisierung gegen Malaria, Vektor besagte DNA-Sequenz enthaltend, Impfstoff das gp190/MSP-1 Oberflächenprotein oder die dafür kodierende DNA enthaltend, Verfahren zur Herstellung des gp190/MSP-1 Oberflächenproteins

2. Ansprüche: 36, teilweise 37-40

Verfahren zur Stabilisierung von Gen-Sequenzen durch Verringerung des AT-Gehaltes der Sequenzen, stabilisiertes Gen, Vektor dieses Gen enthaltend, Impfstoff diesen Vektor enthaltend

THIS PAGE BLANK (USPTO)

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES

IPK 6 C12N15/30 C07K14/445 C12N15/62 A61K39/015 A61K31/70
C07H21/00 C12N1/21 C12N5/10 C12N1/11 C12N15/67

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 6 C12N C07K A61K C07H

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie ^o	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	WO 94 28930 A (VIROGENETICS CORP) 22.Dezember 1994	1-4,12, 14,17, 26-29, 33,34, 39,40 18,20
Y	siehe Seite 19 - Seite 22; Ansprüche 1-33; Beispiele 5,29,53,57,63 ---	
Y	EP 0 340 359 A (WELLCOME FOUND) 8.November 1989 siehe Abbildung 2; Beispiele 3,4 --- -/--	18



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

^o Besondere Kategorien von angegebenen Veröffentlichungen :

- *A* Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
- *E* Älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
- *L* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
- *O* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
- *P* Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

Z Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

6.Mai 1998

Absendedatum des internationalen Recherchenberichts

1 8. 05. 98

Name und Postanschrift der Internationalen Recherchenbehörde
Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Espen, J

THIS PAGE BLANK (USPTO)

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	KASLOW DC ET AL: "Expression and antigenicity of Plasmodium falciparum major merozoite surface protein (MSP1(19)) variants secreted from Saccharomyces cerevisiae." MOL BIOCHEM PARASITOL, FEB 1994, 63 (2) P283-9, NETHERLANDS, XP000603953 siehe das ganze Dokument ---	20
X	HOLDER A A ET AL: "PRIMARY STRUCTURE OF THE PRECURSOR TO THE THREE MAJOR SURFACE ANTIGENS OF PLASMODIUM FALCIPARUM MEROZOITES" NATURE, Bd. 317, 19.September 1985, Seiten 270-273, XP000604859 siehe Abbildung 2 ---	17
X	MYLER P J: "Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plasmodium falciparum Palo Alto PLF-3/B11" NUCLEIC ACIDS RESEARCH., Bd. 17, Nr. 13, 1989, OXFORD GB, Seite 5401 XP002057620 siehe das ganze Dokument ---	17
X	EP 0 154 454 A (WELLCOME FOUND) 11.September 1985 siehe das ganze Dokument ---	1,12,13, 17, 26-28, 39,40
X	SIDDIQUI W A ET AL: "Merozoite surface coat precursor protein completely protects Aotus monkeys against Plasmodium falciparum malaria" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., Bd. 84, 1987, WASHINGTON US, Seiten 3014-3018, XP002057621 in der Anmeldung erwähnt siehe das ganze Dokument ---	33-35, 39,40
X	GENTZ R ET AL: "Major surface antigen p190 of Plasmodium falciparum: detection of common epitopes present in a variety of plasmodia isolates" EMBO JOURNAL., Bd. 7, Nr. 1, 1988, EYNSHAM, OXFORD GB, Seiten 225-230, XP002057622 siehe Abbildung 1; Tabelle 1 ---	17-28, 33,34, 39,40

-/--

THIS PAGE BLANK (USPTO)

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	PAN W ET AL: "A direct and rapid sequencing strategy for the Plasmodium falciparum antigen gene gp190/MSA1." MOL BIOCHEM PARASITOL, JUL 1995, 73 (1-2) P241-4, NETHERLANDS, XP002057623 siehe Seite 1; Tabelle 1 ---	17
X	EP 0 385 962 A (MONSANTO CO) 5.September 1990 siehe Abbildungen 2-4; Beispiele 2,3 ---	36-40
X	EP 0 359 472 A (LUBRIZOL GENETICS INC) 21.März 1990 siehe Seite 9; Abbildung 1; Tabelle 3 siehe Seite 22, Zeile 1 - Zeile 33 -----	36-40

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

Informa patent family members

International Application No

PCT/ 97/05441

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9428930 A	22-12-94	AU 7060294 A EP 0717636 A	03-01-95 26-06-96
EP 0340359 A	08-11-89	AU 620041 B AU 2156988 A CA 1331155 A DE 3882522 A DE 3882522 T DK 478588 A ES 2058294 T IE 62490 B JP 2115464 C JP 2167088 A JP 8013275 B PT 88362 A,B US 5147788 A	13-02-92 09-11-89 02-08-94 26-08-93 03-03-94 07-11-89 01-11-94 08-02-95 06-12-96 27-06-90 14-02-96 30-11-89 15-09-92
EP 0154454 A	11-09-85	AU 592749 B AU 3904685 A DK 79985 A GB 2154592 A IL 74409 A JP 2584733 B JP 61019490 A JP 6189772 A PH 25993 A US 5597708 A	25-01-90 05-09-85 23-08-85 11-09-85 26-08-94 26-02-97 28-01-86 12-07-94 13-01-92 28-01-97
EP 0385962 A	05-09-90	AU 638438 B AU 5163090 A CA 2024811 A EP 0413019 A IL 93513 A JP 3504333 T WO 9010076 A US 5500365 A	01-07-93 26-09-90 25-08-90 20-02-91 14-11-95 26-09-91 07-09-90 19-03-96
EP 0359472 A	21-03-90	AT 132193 T AU 623429 B AU 4118289 A	15-01-96 14-05-92 15-03-90

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/97/05441

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0359472 A		CN 1044298 A	01-08-90
		DE 68925253 D	08-02-96
		EP 0682115 A	15-11-95
		ES 2083384 T	16-04-96
		JP 2186989 A	23-07-90
		US 5380831 A	10-01-95
		US 5567600 A	22-10-96
		US 5567862 A	22-10-96

THIS PAGE BLANK (USPTO)

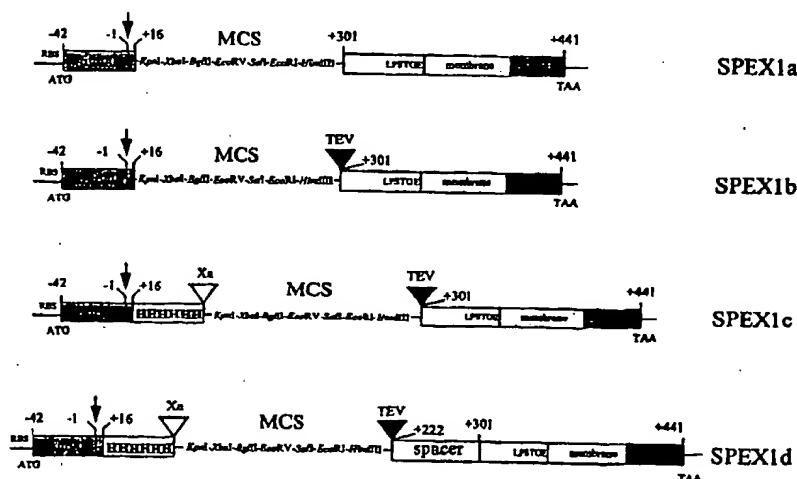
PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/62, 15/31, 15/74, C07K 14/315		A1	(11) International Publication Number: WO 96/40943
(21) International Application Number: PCT/US96/09965		(43) International Publication Date: 19 December 1996 (19.12.96)	
(22) International Filing Date: 6 June 1996 (06.06.96)		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data: 08/472,244 7 June 1995 (07.06.95) US			
(71) Applicant (for all designated States except US): THE ROCKEFELLER UNIVERSITY [US/US]; 1230 York Avenue, New York, NY 10021 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): DARZINS, Aldis [US/US]; 102 South Timber Top, Woodlands, TX 77380 (US). WHITEHEAD, Stephen [US/US]; 7 Prairie Rose Lane, Gaithersburg, MD 20878 (US). HRUBY, Dennis [US/US]; 4017 N.W. Christine, Corvallis, OR 97330 (US).			
(74) Agent: REA, Teresa, Stanek; Burns, Doane, Swecker & Mathis, L.L.P., P.O. Box 1404, Alexandria, VA 22313-1404 (US).			

(54) Title: USE OF GRAM-POSITIVE BACTERIA TO EXPRESS RECOMBINANT PROTEINS



(57) Abstract

A novel system for cloning and expression of genes in gram-positive bacteria. The expression system is based on the finding that many gram-positive bacteria sort proteins to their cell surface through *cis*-acting N-terminal signal sequences and C-terminal anchor regions. In particular, the cell sorting signals of the streptococcal M6 protein, a well-known surface molecule, are used to construct a gram-positive expression system, designated SPEX (Streptococcal Protein Expression). Expression is achieved by cloning the gene of interest into an appropriate SPEX cassette which is then stably introduced into a bacterial host, such as the human commensal *Streptococcus gordonii*. Depending on the SPEX vector used, recombinant proteins can be anchored to the cell wall prior to release by specific endoproteolytic cleavage or secreted into the culture medium during bacterial growth. The use of host bacteria lacking extracellular proteases should protect secreted proteins from proteolytic degradation. Several expression vectors in this system also produce specifically-tagged recombinant proteins which allows for a one-step purification of the resulting product.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

-1-

USE OF GRAM-POSITIVE BACTERIA
TO EXPRESS RECOMBINANT PROTEINS
BACKGROUND OF THE INVENTION

1. *Field of the Invention*

5 The present invention relates to the construction and use of a novel gram-positive expression vector system. Protein fusions containing the amino and carboxy sorting sequences of a gram-positive surface polypeptide can be anchored to the surface of a heterologous host. Alternatively, recombinant proteins may be secreted into the growth medium. This system can be used for
10 overproducing and purifying recombinant proteins or peptides for any purpose, including but not limited to commercial scale production of diagnostic and vaccine antigens and therapeutic proteins, such as hormones and growth factors.

2. *Description of the Related Art*

15 The ability to overproduce many prokaryotic and eukaryotic proteins has been made possible through the use of recombinant DNA technology. The introduction of chimeric DNA molecules into *Escherichia coli* has been the method of choice to express a variety of gene products. The main impetus behind the use of *E. coli*-based protein production systems is the host's short generation time and well-developed genetics. Yet despite the development of
20 many efficient *E. coli*-based gene expression systems in recent years, the most important concern continues to be that associated with downstream processing of the product. Recombinant proteins produced in *E. coli* do not readily cross the outer cell membrane (OM); as a result, polypeptides must be purified from the cytoplasm or periplasmic space (PS). Purification of proteins from these cellular
25 compartments can be somewhat difficult. Frequently encountered problems include low product yields, contamination with potentially toxic cellular material (i.e., endotoxin) and the formation of large amounts of partially folded

-2-

polypeptide chains in non-active aggregates, called inclusion bodies. As a result of these inherent purification difficulties, a great deal of attention has recently been focused on the natural ability of many gram-positive bacteria to export proteins beyond their cell wall boundaries. In spite of the fact that the genetics of gram-positive bacteria are not as well-elucidated as those of *E. coli*, these microorganisms have, nevertheless, become recognized as more favorable candidates as hosts for the production of recombinant proteins.

Proteins exported across the gram-positive cytoplasmic membrane (CM) generally have two fates: they are either released (secreted) into the extracellular milieu or they remain anchored to the cell wall/membrane. Many of the recently developed gram-positive based expression systems have relied exclusively on the former route of protein export (i.e., secretion). The basic strategy for directing the secretion of proteins in gram-positive expression systems involves fusing target proteins with functional N-terminal signal sequences of the gram-positive secretion systems currently available. By far, the most popular hosts belong to the genus *Bacillus*. This group of microorganisms has been used by industry for the production of a variety of economically important proteins (Priest, *Bacteriol. Rev.*, 41:711 (1977); Glenn, *Ann. Rev. Microbiol.*, 30:41 (1976)). Perhaps the most extensively studied gram-positive host-vector expression system is based on the *B. amyloliquefaciens* alpha-amylase gene (Ulmanen *et al.*, *J. Bacteriol.*, 162:176 (1985); Palva *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 79:5582 (1982); Palva *et al.*, *Gene*, 22:229 (1983); Lundstrom *et al.*, *Virus Res.*, 2:69 (1985)). However, expression systems based on other gram-positive hosts such as *B. subtilis*, *B. coagulans*, *B. licheniformis* (U.S. Patents 4,824,782, 5,171,673, 4,711,843; WO patent 8,605,812; Chang, S., *Methods in Enzymology*, 153:507-516 (1987)), *Lactococcus* and *Lactobacillus* spp. (U.S. Patent 5,242,821), *Staphylococcus* spp. (Abrahmsen *et al.*, *EMBO J.* 4:3901-3906 (1986)), *Streptomyces* spp. (U.S. Patent 4,745,056; JP Patent 2,002,379; EP 352,707) and *Corynebacterium* spp. (U.S. Patent 4,965,197; WO Patent 9,303,158) have also been developed for general use.

-3-

Gram-positive secretion systems provide several advantages over traditional *E. coli* based expression systems. One obvious benefit is that proteins exported beyond the cell wall usually retain their native conformation. As a result, one can take full advantage of established purification protocols which are based on the functional properties of the active protein. In addition, gram-positive expression systems usually generate higher protein yields which are generally free of potentially toxic contaminating cellular material.

There are, however, several practical limitations to the purification of extracellular proteins. The first concern is that of protein instability. Recombinant proteins secreted by many gram-positive hosts are extremely sensitive to proteolytic degradation by host-encoded extracellular proteases. The presence of such proteases can drastically affect protein yield. Fortunately, significant improvements have been achieved in maintaining the viability of secreted proteins by genetically modifying host strains with reduced extracellular protease activities (Kawamura *et al.*, *J. Bacteriol.*, 160:442 (1984); U.S. Patent 5,084,383).

Another important point to consider when using gram-positive secretion systems is the requirement for extensive downstream processing. It is widely accepted that the purification of recombinant proteins from bulky, large volume fermentations is an extremely time consuming and costly proposition, both in terms of equipment and manpower. In those cases where the use of a gram-positive secretion system is simply not practical, it is common to resort to a less expensive, gram-negative based expression system. The decision to use *E. coli* as an alternative expression host, despite the potential pitfalls, is made simply on the basis that recombinant polypeptides remain associated with the bacterial cell (i.e., intracellular or periplasmic) and hence are generally easier to purify. Therefore, it is clear that what is needed in this art is a bacterial expression system that incorporates the salient features of both gram-positive (i.e., protein secretion) and gram-negative (i.e., protein compartmentalization) systems. A novel, but simple alternative to the currently available bacterial expression

-4-

systems would be the development of a gram-positive system that is able to specifically anchor or attach recombinant proteins directly to the cell wall surface. In this way the anchoring process can become an integral part of the purification process. For example, cells harboring a recombinant protein
5 attached to the cell surface would be washed, collected, resuspended in a small volume and then treated with a specific agent to affect the release of the desired protein. Following removal of the bacterial cells, the resulting supernatant fluid would be highly enriched for the protein product.

The cell wall of gram-positive bacteria is a complex organelle, which is
10 assembled from peptidoglycans, carbohydrates, and proteins with different biological properties. Surface proteins differ from naturally secreted products in that the former require specific sorting signals that presumably allow them to deviate from the normal default pathway of protein export. Despite the fact that many of the biologically important gram-positive proteins, including the
15 streptococcal M protein, and the protein A and fibronectin binding proteins of *S. aureus*, are anchored at the cell surface, the cell wall of gram-positive bacteria remains a relatively unknown cellular compartment.

While the normal N-terminal export signals that serve as the basis for bacterial secretion systems have been extensively characterized (Abrahmsen *et al.*, *EMBO J.*, 4:3901-3906 (1985); Heijne and Abrahmsen, *FEBS Letters*,
20 244:439-446 (1989); Pugsley, *Microbio. Reviews*, 57:50-108 (1993)), the *cis*-acting sorting signals that orchestrate the anchoring of cellular proteins to the gram-positive cell wall have been only recently elucidated. Sequence alignment of well over 50 gram-positive surface proteins to date (excluding those from *B. subtilis* and *S. pneumoniae*) has revealed the existence of both amino- and
25 carboxy-terminal secretion/anchoring signals. As expected, all of the proteins examined contain an N-terminal signal sequence that presumably directs them to the cellular export machinery. The C-termini of these proteins contain a predominantly hydrophobic, potential membrane spanning region that is
30 immediately followed by a charged tail (Fischetti *et al.*, *Mol. Microbiol.*, 4:1603-

-5-

1605; Figure 1: Darkly shaded region = LP(X)TG(X) consensus sequence;
lightly shaded region = carboxy-terminal hydrophobic, potential membrane
spanning region; residues in bold = charged protein tail; Protein A =
Staphylococcus aureus protein A; M6 = *Streptococcus pyogenes* M6 protein;
5 WapA = *S. mutans* wall-associated protein A; M49 = *S. pyogenes* M49 protein;
IgA-BP = streptococcal IgA-binding protein; Protein G = streptococcal protein
G; Fn-BP = staphylococcal fibronectin-binding protein; T6 = streptococcal T6
protein; Pac = *S. mutans* surface protein; Wg2 = *S. cremoris* surface protease).
Preceding the hydrophobic domain is a proline/glycine-rich region which has
10 been predicted to span the peptidoglycan cell wall in a beta-sheet-like
conformation. Within this region, all of the surface proteins examined have a
nearly 100% conserved hexapeptide with the consensus LP(X)TG(X), where X is
usually a Thr or Ser and sometimes a Gly, Lys, or Asn (Figure 1). The
conservation of the C-terminal sequence elements (i.e., LP(X)TG(X) motif,
15 hydrophobic membrane spanning region and charged tail) suggests that the
process of sorting and anchoring polypeptides to the bacterial cell wall is shared
by many gram-positive species (Fischetti *et al.*, *Mol. Microbiol.*, 4:1603-1605
(1990); Fischetti *et al.*, *Curr. Opinion Biotechnol.*, 4:603-610 (1993)).

The functional importance of the carboxy-terminal sorting elements
20 identified by sequence alignment analysis has been confirmed using protein A
(Moks *et al.*, *Eur. J. Biochem.*, 156:637-643 (1986)), a well-characterized
surface protein of *S. aureus*, as a model system. Protein A, a single polypeptide
chain, contains two major functional domains. The N-terminal domain contains
a 36 amino acid signal peptide sequence followed by several repeated sequence
25 modules (E, D, A, B, and C) that have immunoglobulin binding activity. The
C-terminal domain contains conserved sorting signals which are thought to be
responsible for anchoring protein A to the cell wall (Guss *et al.*, *Eur. J.*
Biochem., 138:413-420 (1984); Guss *et al.*, *Eur. J. Biochem.*, 143:685 (1984);
Fischetti *et al.*, *Mol. Microbiol.*, 4:1603-1605 (1990)). Using deletion analysis,
30 Schneewind *et al.* (*Cell*, 70:267-281 (1992)) confirmed that the proper sorting of

-6-

protein A in *S. aureus* requires all three of the conserved sequence elements (i.e., -LPETGE-motif, C-terminal hydrophobic domain, and charged tail). Based on their results, Schneewind *et al.* (*Cell*, 70:267-281 (1992)) have proposed the following sequence of events in the anchoring of protein A and presumably many other gram-positive proteins to the cell surface. First, the N-terminal signal sequence of protein A directs the polypeptide into the general secretion pathway (GSP) (for reviews, see Pugsley, *Microbio. Reviews*, 57:50-108 (1993); Salmond and Reeves, *TIBS* 18:7-12 (1993)). During the process of membrane translocation, the main function of the C-terminal charged tail is to prevent secretion of the protein into the medium. Translocation of protein A across the membrane presumably results in the recognition of the C-terminal sorting signal. Following recognition, the LPETGE (LPXTGX) motif is cleaved specifically between the threonine (T) and glycine (G) residues and the resulting N-terminal protein fragment is then covalently linked to the *S. aureus* cell wall (Navarre and Schneewind, *Mol. Microbiol.*, 14:115-121 (1994)).

Because the anchor motif of gram-positive surface proteins is conserved among a wide variety of molecules and several different gram-positive species, a logical question to ask is whether it would be possible to anchor a well-known surface protein to the cell wall of a heterologous gram-positive host. Pozzi *et al.* (*Res. Microbiol.*, 143:449-457 (1992)), using the fibrillar streptococcal M protein, were the first to address the question of protein anchoring in heterologous gram-positive hosts. Structurally, the M protein consists of an extended central alpha-helical coiled-coil rod flanked by functional end domains (Figure 2) (Phillips *et al.*, *Proc. Natl. Acad. Sci. (USA)*, 78:4689-4693 (1981); Fischetti, *Clin. Microbiol. Reviews*, 2:285-314 (1989))(Figure 2, (A) Structural representation of the M6 protein; (B) Linear representation of the M6 protein molecule. The map displays the N-terminal signal peptide (S), several repeated domains (A, B, C, and D), the proline/glycine rich region that contains the LPSTGE motif (LPXTGX), the C-terminal hydrophobic, potential membrane spanning domain (black bar), and charged tail (KRKEEN). Amino and carboxy

-7-

terminal residues important for export and cell wall sorting/anchoring of the M6 protein are shaded. The anchor region is not drawn to scale.). The N-terminus contains a 42-residue signal sequence required for proper processing (Fischetti, *Clin. Microbiol. Reviews*, 2:285-314 (1989); Hollingshead *et al.*, *J. Biol. Chem.*, 261:1677-1686 (1986)), while the C-terminus, as determined by sequence comparisons, contains the conserved residues that are necessary for cell sorting and anchoring (Fischetti *et al.*, *Mol. Microbiol.*, 4:1603-1605 (1990); Pancholi and Fischetti, *J. Bacteriol.*, 170:2618-2624 (1988); Figure 1). After integration of a promoterless *emm-6.1* gene in the chromosome of the human oral commensal *S. gordonii*, it was demonstrated that the M6 protein was correctly localized to the cell wall of the heterologous host (Pozzi *et al.*, *Res. Microbiol.*, 143:449-457 (1992)).

The next logical step was to examine the possibility of exploiting the M6 sorting signals in order to express recombinant proteins on the surface of a gram-positive host. The M6 protein has been successfully modified to deliver several heterologous antigens, such as sequences derived from the human papilloma virus type 16 E7 protein, the immunodominant epitope of HIV-1 gp120 and the major allergen of white-face hornet venom, to the surface of *S. gordonii* (Pozzi *et al.*, *Infect. Immun.*, 60:1902-1907 (1992); Pozzi *et al.*, *Vaccine*, 12:1071-1077 (1994); Medaglini *et al.*, *Proc. Natl. Acad. Sci. (USA)*, in press (1995)). Gram-positive oral commensal bacteria expressing recombinant fusion proteins on their cell surface have been used to elicit both a mucosal and systemic immune response to foreign antigens while colonizing the oropharynx (Fischetti *et al.*, *Curr. Opinion Biotechnol.*, 4:603-610 (1993); Medaglini *et al.*, *Proc. Natl. Acad. Sci. (USA)*, in press (1995)).

The recent results of the M6 experiments are important for two main reasons. First, they serve to confirm the hypothesis that the mechanism of protein sorting and anchoring is an extremely well conserved, if not universal, process in gram-positive bacteria. More importantly, these experiments are evidence for proof of the concept that it is possible to exploit cell wall protein

-8-

sorting signals for the purpose of constructing novel expression systems that anchor recombinant proteins to the surface of a gram-positive bacterial host. Therefore, in view of the aforementioned deficiencies attendant with prior art expression systems, it should be apparent that there still exists a need in the art
5 for the construction and use of a unique gram-positive protein expression system.

SUMMARY OF THE INVENTION

Accordingly, a major objective of the present invention is to develop a gram-positive bacterial expression system that can be used for the production and purification of recombinant proteins.

10 It is also the objective of this invention to provide, as a specific example, the construction and use of the SPEX system (for Streptococcal protein expression), which is based on the streptococcal M6 protein. This system will allow one familiar in the art of genetic manipulation and protein biochemistry to express and purify recombinant proteins from gram-positive hosts, such as the
15 human oral commensal *S. gordonii*.

More specifically, the present invention provides a method for expressing proteins in gram-positive bacteria, comprising selecting a heterologous bacterial host which is gram-positive, deficient in extracellular protease production, and capable of anchoring surface proteins via C-terminal sorting signals; introducing
20 into the host a plasmid vector comprising DNA fragments encoding amino and carboxy terminal sorting sequences of a gram-positive cell wall surface protein, DNA fragments coding for the desired protein product, and transcriptional and translational control sequences that function in the chosen host; expressing a fusion protein comprising the desired protein product, an N-terminal signal
25 sequence, and a C-terminal sorting sequence wherein the expressed protein is anchored to the host's cell surface; cleaving the anchored protein from the host's cell surface using a protease; and purifying the cleaved protein from the supernatant fluid. The present invention also provides an alternative method for expressing proteins in a gram-positive host whereby the desired protein product

-9-

is secreted into the extracellular space, and for purifying the secreted proteins from the supernatant fluid.

Briefly, the present invention features general methods for the construction and use of a gram-positive host/vector system that uses the cell sorting signals of surface proteins for targeted attachment to the cell wall. To demonstrate the utility of this system, the invention also describes the construction of several specialized vectors, designated SPEX, which are based on the *S. pyogenes* M6 protein. These vectors are intended to be used for the expression of proteins and peptides in gram-positive bacteria such as the human oral commensal, *S. gordonii*. A variety of recombinant protein products can be expressed by selecting the appropriate SPEX cloning vector. Several protein configurations are possible such as the presence of a combination of polyhistidine tag and protein A IgG binding domains, factor Xa or TEV protease cleavage sites, and a cell wall anchoring domain LP(X)TG(X). Each of the cloning vectors acts as a shuttle plasmid which permits the initial cloning to be performed in *E. coli*, followed by transformation into a gram-positive host, such as the naturally competent *S. gordonii* V288 (Challis). Depending on the SPEX vector used, recombinant proteins can be purified from an anchored state after cleavage with a specific protease or, alternatively, from the culture medium.

With the foregoing and other objectives, advantages and features of the invention that will become hereinafter apparent, the nature of the invention may be more clearly understood by reference to the following detailed description of the preferred embodiments of the invention and to the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the alignment of the C-terminal residues from a select group of gram-positive surface proteins.

Figure 2 depicts the functional domains of the *S. pyogenes* M6 protein.

Figure 3 depicts the nucleotide sequence of the *S. pyogenes emm6* gene

-10-

and adjacent DNA regions.

Figure 4 is a diagrammatic representation of the construction of SPEX vectors. The pBLUESCRIPT (Ap^r) backbone of each plasmid is not shown.

Figure 5 illustrates possible SPEX cassettes useful for anchored surface production and purification of recombinant proteins.

Figure 6 illustrates possible SPEX cassettes used for the secretion and purification of recombinant proteins.

Figure 7 depicts some possible *S. gordonii* strains for integrated expression of M6 fusion proteins.

Figure 8 depicts integration of the M6 fusions into the *S. gordonii* chromosome.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS OF THE INVENTION

More particularly, the present invention relates to general methods for the construction and use of a novel gram-positive expression system. Proteins produced by the expression system described herein can be purified either directly from the cell surface or from the supernatant fluid following secretion. For example, an expression system based on the well-known M6 protein of *S. pyogenes* is also described. This expression system, designated SPEX (for Streptococcal protein expression), uses the human oral commensal *Streptococcus gordonii* as a model host for this group of organisms.

A. General Method

The first major factor to be considered when constructing the gram-positive expression system is to determine which bacterium to use as a host. To be effective in the present invention the bacterial host should be a gram-positive bacteria that is capable of anchoring surface proteins via C-terminal sorting signals. Such gram-positive bacteria include, but are not limited to, the genera

-11-

Aerococcus, *Bifidobacterium*, *Corprococcus*, *Deinobacter*, *Deinococcus*,
Enterococcus, *Gemella*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*,
Marinococcus, *Melissococcus*, *Micrococcus*, *Pediococcus*, *Peptococcus*,
Peptostreptococcus, *Planococcus*, *Ruminococcus*, *Saccharococcus*, *Salinococcus*,
5. *Carcina*, *Staphylococcus*, *Stomatococcus*, *Streptococcus*, *Trichococcus*, and
Vagococcus. To gain the maximum benefit from the present invention, the
bacterial host should be deficient in extracellular protease production.

In the present invention the more preferred species to be used as an
expression host is *Streptococcus gordonii*. The most preferred strain is a
10 protease deficient variant of V288 (Challis).

The second major consideration is the selection of the gram-positive cell
wall protein to be used as the basis for the construction of the expression vectors.
Nearly all surface molecules of gram-positive bacteria have amino and carboxy
terminal sorting signals that specifically target them for cell wall attachment
15 (Fischetti *et al.*, *Mol. Microbiol.*, 4:1603-1605 (1990)). The expression system
used in the practice of the present invention can be constructed with any gram-
positive surface protein that contains the appropriate sorting signals. Such
sorting signals can be readily identified by those skilled in the art. Sources of
surface proteins include, but are not limited to, such gram-positive genera as
20 *Aerococcus*, *Bifidobacterium*, *Corprococcus*, *Deinobacter*, *Deinococcus*,
Enterococcus, *Gemella*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Marinococcus*,
Melissococcus, *Micrococcus*, *Pediococcus*, *Peptococcus*, *Peptostreptococcus*,
Planococcus, *Ruminococcus*, *Saccharococcus*, *Salinococcus*, *Carcina*,
Staphylococcus, *Stomatococcus*, *Streptococcus*, *Trichococcus*, and *Vagococcus*.

25 The more preferred protein to be used in the present invention is the type-
6 M protein of *Streptococcus pyogenes*. The structural gene for the type 6-M
protein from *S. pyogenes* has been cloned (Scott and Fischetti, *Science*, 221:758-
760 (1983)) and the complete nucleotide sequence has been determined
(Hollingshead *et al.*, *J. Biol. Chem.* 261:1677-1686 (1986)). This sequence is
30 shown in Figure 3. The predicted M6 amino acid residues are shown below the

-12-

DNA sequence. The putative ribosome binding site (RBS) is boxed and the signal peptide (-42 to -1) is underlined. The mature polypeptide represents residues +1 to 441. The proline/glycine rich cell wall domain is shown as unshaded rectangles. The LPXTGX (LPSTGE) motif is shown in bold letters.

5 The adjacent lightly shaded rectangle represents the hydrophobic, membrane spanning domain and the darkly shaded rectangle represents the C-terminal charged tail. Relevant restriction sites are shown above the DNA sequence. A transcriptional terminator is shown as the inverted repeat structure denoted by the arrows.

10 The next consideration is to isolate and characterize the DNA sequence that codes for the cell wall protein in the chosen host. Methods used for the construction and screening of genomic libraries (i.e., plasmid, cosmid, phage or YAC) to isolate the appropriate protein coding sequence and techniques for determining its nucleotide sequence are known to those skilled in the art of
15 genetic engineering. See, e.g., Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition (1989); Maniatis, Fritsch, & Sambrook, *Molecular Cloning: A Laboratory Manual* (1982) and Griffin & Griffin, *PCR Technology: Current Innovations* (1994).

The DNA fragments that encode the amino and carboxy terminal sorting
20 signals of the surface protein can be isolated using recombinant DNA techniques familiar to those skilled in the art. These fragments can be assembled on an *E. coli* cloning vector such as pUC18/19 or any other appropriate plasmid element. The N-terminal signal should include a complete leader sequence as well as a number of amino acid residues from the N-terminus of the mature
25 protein. The inclusion of these sequences will insure proper targeting to the cellular export machinery and correct processing during the actual translocation process. The C-terminal sorting (i.e., anchor) signals should include the LPXTGX motif (where X is any amino acid residue), a hydrophobic membrane spanning domain and a charged tail. A comparison of such sorting sequences is
30 shown in Figure 1. The darkly shaded sequences on the left side of the

-13-

alignment represents the LP(X)TG(X) consensus sequence and the proteins were aligned from this sequence. The lightly shaded region on the right side of the alignment represents the carboxyl-terminal hydrophobic, potential membrane spanning region. Residues shown in bold represent the charged protein tail.

5 To facilitate the construction of heterologous fusion proteins, a small DNA fragment containing the recognition sequences for several restriction endonucleases can be placed between the DNA sequences that encode the amino and carboxy terminal sorting signals. This multiple cloning site (MCS) will facilitate the insertion of heterologous protein coding sequences and upon
10 expression will generate in-frame protein fusions that are anchored to the gram-positive cell wall surface. Alternatively, expression vectors can also be constructed that lack the C-terminal anchor signals. In this case, the protein fusions will not be targeted for anchoring but, instead, will be secreted into the extracellular space.

15 The vectors described in the present invention will allow one familiar in the art to produce recombinant proteins that are either anchored at the cell surface or secreted into the extracellular space. Regardless of the final destination of the target protein, all of the vectors contain the M6 ribosome binding site (RBS), translation initiation codon (ATG) and the 42 amino acid
20 leader sequence (-42 to -1) (Figure 3, 4 and 5). In addition, all of the vectors contain some portion of the mature N-terminus of the M6 protein. Previous studies have demonstrated that as many as 122 amino acids of the mature M protein N-terminus are also required for successful translocation. However, more recently, it has been shown that as few as 5 N-terminal amino acid residues
25 of the mature protein may be required for proper processing and translocation of a M6 fusion protein to the cell surface (Fischetti *et al.*, *Curr. Opinion Biotechnol.* 4:603-610 (1993)). The preferred number of M6 amino terminal residues to be used in the practice of the present invention is 58 and include residues -42 to +16 (Figure 3).

30 The preferred C-terminal sorting signals used in the present invention are

-14-

also derived from the M protein of type-6 *S. pyogenes*. The vectors designed for targeted protein anchoring contain the M6 C-terminal residues 302 to 441, termination codon (TAA) and downstream sequences, which include a putative transcriptional terminator (Figure 3). Vectors that are designed for protein secretion do not contain the M6 C-terminal sorting signals.

The expression vector systems of the present invention are suitable for overproducing and purifying any desired protein. At the very least, the expression vector systems of the present invention are useful for overproducing and purifying recombinant proteins or peptides. These proteins or peptides may be used for any purpose; for example, they may be used to generate diagnostic and vaccine antigens and therapeutic proteins, such as hormones and growth factors, and the like.

To facilitate the release of anchored fusion proteins from the gram-positive cell wall, cleavage sites for various proteolytic enzymes can be engineered into the expression vectors. These unique sequences can optionally be incorporated into the vectors so that they reside either immediately downstream of the N-terminal signal peptide or immediately upstream of the C-terminal sorting signal domain. It is also possible to engineer cleavage sites at both locations. Proteases that may be suitable include, but are not limited to, the proteinase of tobacco etch virus (TEV), thrombin, enterokinase, and factor Xa.

A variably sized protein spacer or linker region that physically displaces the fusion protein away from the cell surface may also be incorporated into the expression vectors. It is possible that in the absence of these spacers the efficiency of protease cleavage may be reduced due to steric hinderance with the cell wall surface. These linker sequences can be incorporated between the protease recognition sequence and the C-terminal sorting signal domain.

Any unique protease cleavage site can be incorporated into the vectors of the present invention. Preferably, the protease cleavage site will be greater than 4 amino acid residues. The most preferred protease for use with the present invention is the TEV NIa proteinase. The TEV NIa proteinase cleaves a specific

-15-

consensus cleavage site which spans the seven amino acid sequence E-X-V/I/L-Y-X-Q*S/G (X can be any amino acid residue; Dougherty *et al.*, *EMBO J.*, 7:1281-1287 (1988)). The preferred cleavage site for the present invention is E-N-L-Y-F-Q*G (Parks *et al.*, *Anal. Biochem.* 216:413-417 (1994)). TEV

5 cleavage sites can be engineered anywhere along the fusion construct; however, the preferred location is immediately adjacent to the proline/glycine rich region harboring the LPXTGX motif (see Figure 5: RBS = M6 ribosome binding site; ATG = initiation codon; MCS = multiple cloning site (MCS); black bar = membrane spanning domain; closed triangles = engineered cleavage site for

10 TEV (tobacco etch virus) NIa protease; open triangle = cleavage site for factor Xa (IEGR/); small vertical arrow 16 = putative cellular signal peptidase cleavage site; see SPEX1b and SPEX1c). The addition of a protein spacer region between the TEV recognition sequence and the glycine/proline-rich anchor region may improve TEV NIa proteinase efficiency. The preferred spacer for the

15 present invention is derived from the M6 protein and contains amino acid residues 222 to 301 (Figure 5; SPEX1d).

Following protease treatment, the released proteins can be purified in a variety of ways known in the art. Suitable purification methods include, but are not limited to, dialysis, ultrafiltration, zonal centrifugation, molecular-exclusion

20 chromatography, isoelectric precipitation, solvent fractionation, electrophoretic separation, thermal precipitation, ion-exchange chromatography, affinity chromatography, and the like. A single step purification using affinity tags is preferred. These affinity tags, like the protease cleavage sites, can be optionally engineered at either the amino or carboxy terminal regions of the fusion protein.

25 Useful affinity tags include, but are not limited to, a polyhistidine tract (HHHHHH), the IgG binding domain of protein A and glutathione S-transferase (GST). Recombinant proteins released from the surface of a gram-positive host can be easily purified in a one step process using metal chelation (e.g., Ni-agarose), protein A-Sepharose, and glutathione-Sepharose column

30 chromatography. N-terminal signal and affinity tag sequences can be easily

-16-

removed by incorporating a protease cleavage site that is different from the recognition sequence used to remove the protein from the cell surface.

The preferred affinity tag for the present invention is a consecutive stretch of 6 to 10 histidine residues (HHHHHH). A polyhistidine tag of six amino acid residues has been shown to be poorly immunogenic and rarely affects protein function and structure. The polyhistidine affinity tag can be engineered at either the amino or carboxy terminus of the protein. For the present invention, the preferred site is immediately downstream of the N-terminal M6 sequences (Figure 5).

The N-terminal M6 and affinity tag sequences can be removed from the purified proteins by engineering in a second protease cleavage site. The preferred enzyme for the present invention is factor Xa. The preferred location is immediately downstream of the affinity tag (Figure 5).

In order to efficiently express the recombinant gene which encodes the fusion protein, regulatory sequences should preferably be included that assure adequate transcription and translation. These vectors should contain transcriptional (i.e., promoters) and translational (i.e., ribosome binding site) control sequences that function in the chosen gram-positive host.

A number of promoters may optionally be used in the practice of the present invention. These regulatory sequences, which are known to those skilled in the art, include heterologous promoters from both gram-negative and gram-positive bacteria. These regulatory sequences can cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus. Alternatively, constitutively expressed promoter elements can also be used.

Gene dosage is a variable that has been shown to possibly have an effect on protein production in several bacterial expression systems. In general, genes present on multiple copy plasmids generate higher levels of protein. The plasmids used for the expression of recombinant proteins in the present invention should be capable of stably replicating in the chosen gram-positive host. These plasmids, which replicate at either a high or low copy number, are familiar to

-17-

those skilled in the art.

In the present invention, the extrachromosomal elements used in plasmid-linked expression of recombinant genes may optionally contain appropriate genetic markers for selection purposes. Suitable genetic markers include various antibiotic and heavy metal resistance genes as well as nutritional determinants. However, other genes that provide a selective advantage can also be used. Ideally the selective markers should be capable of being expressed in both *E. coli* and gram-positive hosts. This would allow one to carry out the initial cloning in *E. coli* with subsequent transfer of the recombinant plasmid to a gram-positive host. Furthermore, in order to facilitate the insertion of heterologous DNA fragments, plasmids should also contain an in-frame polylinker or MCS which includes useful restriction endonuclease cleavage sites.

Techniques for introducing recombinant plasmids into nontransformable hosts are familiar to those skilled in the art. These methods include, but are not limited to, electroporation (Dower, in *Genetic Engineering-Principles and Methods* Vol.12, pp.275-296 (1990)), chemical transformation, conjugation, transduction, and the like. Recombinant DNAs can be easily introduced into those gram-positive species that are naturally competent by transformation.

The presence of certain genes on multiple copy plasmids can occasionally cause toxicity and, in some cases, cell death. The toxic effects of these gene products can many times be reduced by simply decreasing the copy number of the determinant to be expressed. This can be readily achieved by stably integrating the gene encoding the fusion product directly into the host chromosome. The genetic methods that are commonly used to create single copy prokaryotic expression modules involve homologous, site specific and illegitimate (transposition) recombination.

The preferred method of gene expression for the present invention involves integration of the recombinant SPEX construct into the *S. gordonii* chromosome. *S. gordonii* strains that are useful as expression hosts include GP230 (Pozzi *et al.*, *Res. Microbiol.* 143:449-457 (1992)), GP232 (Pozzi *et al.*,

-18-

Infect. Immun., 60:1902-1907 (1992)) and GP251 (Pozzi *et al.*, *Vaccine*, 12:1071-1077 (1994); Figure 7: *emm-6* is the structural gene for the *S. pyogenes* type 6 M protein. The *ermC* and *cat* genes encode erythromycin and chloramphenicol resistance, respectively. Thin lines flanking the inserted (boxed) sequences represents the *S. gordonii* chromosome. The integrated sequences are located downstream of a strong chromosomal *S. gordonii* promoter (P)). These strains contain genetic cassettes that have been integrated downstream of a strong *S. gordonii* promoter via homologous recombination. In the present invention, these cassettes provide flanking regions of homology that will allow one skilled in the art to replace the sequences present on the GP230, GP232 and GP251 with the recombinant SPEX plasmids constructed *in vivo*.

The following examples are presented in order to more fully illustrate the preferred embodiments of the invention. They should in no way be construed, however, as limiting the broad scope of the invention.

EXAMPLE 1: Construction of SPEX plasmids

A. Anchored vectors

Plasmid pVMB20 (Pozzi *et al.*, *Infect. Immun.* 60:1902-1907 (1992)) was constructed by subcloning a 3.4 kb *ClaI* fragment, which contains a promoterless *emm-6* gene and a gene encoding erythromycin resistance (*ermC*; Horinouchi and Weisblum., *J. Bacteriol.* 150:804-814 (1982)), from pVMB3 (Pozzi *et al.*, *Res. Microbiol.* 143:449-457 (1992)) into the plasmid pBLUESCRIPT. A 545 bp interior region of the *emm-6* gene was removed by *KpnI* and *HinDIII* digestion and replaced with a small DNA fragment harboring a synthetic multiple cloning site (MCS). The resulting plasmid, designated pSMB55 (Figure 4), contains sequences encoding the first 122 N-terminal and the last 140 C-terminal amino acids of the M6 protein. DNA fragments encoding heterologous protein sequences are inserted into the MCS of pSPEX1a or pSPEX2a to form in-frame fusions with the M6 N-terminal and/or C-terminal sequences.

-19-

The *SacI-HpaI* fragment from within the *ermC* determinant of pSMB55 was removed and replaced with a fragment harboring the *aphIII* gene encoding kanamycin resistance (Trieu-Cuot and Courvalin, *Gene*, **23**:331-341 (1983)). The resulting plasmid was designated pSMB113 (Figure 4).

5 Plasmid pSMB104 carries the *emm-6*_{Δ104} gene encoding for M6_{Δ104}, a protein containing the first 16 N-terminal and the last 220 C-terminal amino acids (Pozzi, G., unpublished data). The *ClaI-KpnI* fragment of pSMB104 was used to replace the existing *ClaI-KpnI* fragment of pSMB113. The resulting plasmid vector, designated pSPEX1a (Figure 4), can be used to construct protein fusions
10 that contain the N-terminal 58 and C-terminal 140 amino acids of the M6 protein. These protein fusions, when expressed in an appropriate gram-positive host, will be anchored to the bacterial cell surface (Figure 5).

B. Secretion vectors

Deletion of the *HinDIII-SacI* fragment of pSPEX1a effectively removes
15 the C-terminal sorting signals of the M6 protein. The resulting plasmid, designated pSPEX2a, retains the M6 N-terminal 58 amino acid residues necessary for protein translocation and processing (Figure 4). As a result, protein fusions generated with pSPEX2a will be secreted into the growth medium. TEV proteinase cleavage sites and polyhistidine tags can be
20 incorporated to facilitate the removal of non-essential protein sequences and for purification purposes, respectively (Figure 6: RBS = M6 ribosome binding site; ATG = initiation codon; Term = termination codon; triangle = engineered cleavage site for TEV (tobacco etch virus) NIa proteinase; shaded residues = important for proper processing and secretion of M6; small vertical arrow =
25 putative cellular signal peptidase cleavage site).

-20-

EXAMPLE 2. Expression of recombinant genes in *S. gordonii*

In the present invention, recombinant genes encoding M6 fusion proteins can be expressed from multicopy plasmids. Since the preferred expression host in the present invention is *S. gordonii*, plasmids containing an origin of replication (*oriV*) that functions in a variety of streptococcal species can be used (Horodniceanu *et al.*, *Antimicrob. Agents Chemother.*, 10:795-801 (1976); Clewell *et al.*, *J. Bacteriol.*, 117:283-289 (1974); Behnke and Ferretti, *Plasmid*, 4:130-138 (1980); Macrina *et al.*, *Infect. Immun.* 28:692-699 (1980); Macrina *et al.*, *Gene* 19:345-353 (1982)). Alternatively, recombinant genes can be expressed in the form of single copies which have been stably inserted into the *S. gordonii* chromosome.

The recombinant plasmid is naturally linearized during transformation (Pozzi *et al.*, *Res. Microbiol.*, 141:659-670 (1990); Pozzi *et al.*, *Infect. Immun.*, 60:1902-1907 (1992)) and recognition of the flanking homologous segments facilitates integration of the M6 gene fusion, together with the *aphIII* gene (Figure 8: the *cat* gene present in the *S. gordonii* recipient strain GP251 confers chloramphenicol resistance). Recombinant *S. gordonii* are selected for kanamycin resistance (conferred by *aphIII*) using the "multilayer" plating technique as previously described by Pozzi *et al.* (*FEMS Microbiol. Lett.*, 48: 189-194 (1987)). Kanamycin-resistant transformants are then scored for loss of either erythromycin (conferred by *ermC*) or chloramphenicol (conferred by *cat*) resistance (depending on the recipient). Antibiotics are used at the following concentrations: erythromycin, 5 µg/ml, chloramphenicol, 5 µg/ml, and kanamycin, 500 µg/ml. Southern blot (Southern, *J. Mol. Biol.*, 98:503 (1975); Southern, *Methods Enzymol.*, 69:152 (1980)) or PCR analysis of purified genomic DNA of Km^r(Cm^s or Em^s) transformants can be used to establish the structure and copy number of the expression cassette in the bacterial chromosome.

Production of the recombinant protein can be determined by Western blot

analysis of whole cell extracts (Towbin *et al.*, *Proc. Natl. Acad. Sci. (USA)*, **76**:4350 (1979); Towbin and Gordon, *J. Immunol.*, **72**:313 (1984)). Surface expression of the M6 fusion protein can be tested by immunofluorescence according to the procedure described previously (Pozzi *et al.*, *Vaccine*, **12**:1071-1077 (1994)). Recombinant proteins can be detected using polyclonal antisera against the entire M6 molecule, monoclonal antibodies directed against specific M6 epitopes (Jones *et al.*, *J. Exp. Med.*, **164**:1226-1238 (1986); Jones and Fischetti, *J. Exp. Med.*, **167**:1114-1123 (1988)) or antisera directed against the heterologous protein expressed as a part of the fusion.

10

15

20

25

-22-

(Qiagen Inc.). Following washing, the recombinant protein is then eluted from the column using an imidazole gradient under mild conditions (i.e., pH 6.0). The amount of purified protein can be measured by the method of Bradford (Bradford, *Anal. Biochem.*, 72:248-254 (1976)).

- 5 The use of the SPEX2 series of vectors (Figure 6) generates fusion proteins that are secreted into the surrounding medium during growth. Following removal of the bacterial cells, the spent growth medium is retained and used as the starting material for purification of the desired product. Recombinant proteins can be purified using metal chelation chromatography, as
10 described above, and non-essential N-terminal residues can be removed with TEV N1a proteinase treatment.

- While the invention has been described and illustrated herein by references to various specific materials, procedures, and examples, it is understood that the invention is not restricted to the particular combinations of
15 materials and procedures selected for that purpose. Numerous variations of such details can be implied and will be appreciated by those skilled in the art.

-23-

WHAT IS CLAIMED IS:

1. A method for expressing a desired protein in gram-positive bacteria,
comprising:
 - selecting a heterologous bacterial host which is
5 gram-positive, and
 capable of anchoring surface proteins via C-terminal sorting signals;
 introducing into the host a plasmid vector comprising
 DNA fragments encoding
 amino and carboxy terminal sorting sequences of a gram-positive
10 cell wall surface protein and
 the desired protein, and
 transcriptional and translational control sequences that function in the
 chosen host;
 expressing a fusion protein comprising
15 the desired protein,
 an N-terminal signal sequence, and
 a C-terminal sorting sequence
 wherein the expressed fusion protein is anchored to the host's cell
 surface;
20 cleaving the anchored protein from the host's cell surface using a
 protease; and
 recovering and purifying the cleaved protein.
2. The method of claim 1 wherein the host is selected from the genus
25 *Aerococcus*, *Bifidobacterium*, *Corprococcus*, *Deinobacter*, *Deinococcus*,
 Enterococcus, *Gemella*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*,
 Marinococcus, *Melissococcus*, *Micrococcus*, *Pediococcus*, *Peptococcus*,
 Peptostreptococcus, *Planococcus*, *Ruminococcus*, *Saccharococcus*,
 Salinococcus, *Carcina*, *Staphylococcus*, *Stomatococcus*, *Streptococcus*,
 Trichococcus, or *Vagococcus*.

-24-

3. The method of claim 1 wherein the host is from the genus *Streptococcus*.
 4. The method of claim 1 wherein the host is *Streptococcus gordonii*.
 5. The method of claim 1 wherein the host is *Streptococcus gordonii* V288 (Challis), GP230, GP232, or GP251.
 - 5 6. The method of claim 1 wherein the host is deficient in extracellular protease production.
 7. The method of claim 1 wherein the gram-positive cell wall surface protein is derived from the genus *Aerococcus*, *Bifidobacterium*, *Corprococcus*,
10 *Deinobacter*, *Deinococcus*, *Enterococcus*, *Gemella*, *Lactobacillus*,
Lactococcus, *Leuconostoc*, *Marinococcus*, *Melissococcus*, *Micrococcus*,
Pediococcus, *Peptococcus*, *Peptostreptococcus*, *Planococcus*, *Ruminococcus*,
Saccharococcus, *Salinococcus*, *Carcina*, *Staphylococcus*, *Stomatococcus*,
Streptococcus, *Trichococcus*, or *Vagococcus*.
 - 15 8. The method of claim 7 wherein the gram-positive cell wall surface protein is the type-6 M protein of *Streptococcus pyogenes*.
-
9. The method of claim 7 wherein the gram-positive cell wall surface protein comprises:
 - 20 an N-terminal signal sequence comprising a complete leader sequence and a number of amino acid residues from the N-terminus of the mature protein;
 - a C-terminal sorting signal comprising the LPXTGX motif, wherein X is any amino acid residue, residing within a proline/glycine rich region;
 - a hydrophobic membrane spanning domain; and
 - a charged tail.

-25-

10. The method of claim 9, wherein the gram-positive cell wall surface protein additionally comprises a DNA fragment having a recognition sequence for at least one restriction endonuclease (the multiple cloning site, MCS) located between the DNA sequences that encode the amino- and carboxy-terminal sorting signals.
5
11. The method of claim 1 wherein the plasmid vector further comprises at least one cleavage site for a proteolytic enzyme.
12. The method of claim 1 wherein the protease is the proteinase of tobacco etch virus (TEV), thrombin, enterokinase, or factor Xa.
- 10 13. The method of claim 9 wherein the plasmid vector additionally comprises a variably sized protein spacer or linker sequence that physically displaces the fusion protein away from the cell surface.
14. The method of claim 13 wherein the linker sequence is incorporated between the protease recognition sequence and the C-terminal sorting signal domain.
- 15 15. The method of claim 1 wherein the cleaved proteins are purified in a single step through the use of affinity tags.
16. The method of claim 15 wherein the affinity tags are polyhistidine tracts containing from 6 to 10 histidine residues, the IgG binding domain of protein A, or glutathione S-transferase.
- 20 17. The method of claim 1 wherein the cleaved proteins are purified using metal chelation, protein A-Sepharose, glutathione-Sepharose column chromatography, or a combination thereof.

-26-

18. The method of claim 1 wherein the plasmid vector additionally comprises extrachromosomal genetic markers that are sufficient to provide a selective advantage to a host expressing said genetic markers.
19. The method of claim 18 wherein the genetic markers are antibiotic genes,
5 heavy metal resistance genes, and nutritional determinants.
20. The method of claim 18 wherein the genetic markers are capable of being expressed in both *E. coli* and gram-positive hosts.
21. The method of claim 18 wherein the plasmid vector additionally comprises a
10 polylinker or multiple cloning site which includes restriction endonuclease cleavage sites.
22. The method of claim 7 wherein the plasmid vector comprises sequences coding for the M6 ribosome binding site, the translation initiation codon (ATG), the 42 amino acid leader sequence (-42 to -1), and a portion of the mature N-terminus which is at least 5 amino acids long.
- 15 23. The method of claim 22 wherein the portion of the N-terminus coded for comprises residues -42 to +16.
24. The method of claim 1 wherein the C-terminal sorting signal is derived from the M protein of type-6 *S. pyogenes*.
25. The method of claim 1 wherein the plasmid vector additionally comprises the
20 M6 C-terminal residues 302 to 441, termination codon (TAA), and downstream sequences which comprise a transcriptional terminator.
26. The method of claim 1, wherein the plasmid vector is pSPEX1a.

-27-

27. The method of claim 26, wherein the pSPEX1a vector additionally comprises a protease which has a cleavage site of greater than 4 amino acid residues.
28. The method of claim 26, wherein the protease is the TEV NIa proteinase, which cleaves a specific consensus cleavage site which spans the seven amino acid sequence E-X-V/I/L-Y-X-Q*S/G, wherein X can be any amino acid residue.
29. The method of claim 28, wherein the cleavage site is E-N-L-Y-F-Q*G.
30. The method of claim 28, wherein the cleavage site is located adjacent to the proline/glycine-rich region harboring the LPXTGX motif.
31. The method of claim 28, wherein a protein spacer is added between the TEV proteinase recognition sequence and the proline/glycine-rich anchor region.
32. The method of claim 31, wherein the spacer is derived from the M6 protein and consists of amino acid residues 222 to 301.
33. A method for expressing a desired protein in gram-positive bacteria, comprising:
- selecting a heterologous, gram-positive bacterial host;
 - introducing into the host a plasmid vector comprising DNA fragments encoding
 - amino and carboxy terminal sorting sequences of a gram-positive cell wall surface protein, and
 - the desired protein; and
 - transcriptional and translational control sequences that function in the chosen host;
 - expressing a fusion protein comprising

-28-

the desired protein, and
an N-terminal signal sequence

wherein the expressed protein is secreted from the cell; and
purifying the protein product from the supernatant fluid following
secretion.

- 5
34. The method of claim 33 wherein the host is selected from the genus *Aerococcus*, *Bifidobacterium*, *Corprococcus*, *Deinobacter*, *Deinococcus*, *Enterococcus*, *Gemella*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Marinococcus*, *Melissococcus*, *Micrococcus*, *Pediococcus*, *Peptococcus*,
10 *Peptostreptococcus*, *Planococcus*, *Ruminococcus*, *Saccharococcus*, *Salinococcus*, *Carcina*, *Staphylococcus*, *Stomatococcus*, *Streptococcus*, *Trichococcus*, or *Vagococcus*.
35. The method of claim 33 wherein the host is from the genus *Streptococcus*.
36. The method of claim 33 wherein the host is *Streptococcus gordonii*.
- 15 37. The method of claim 33 wherein the host is *Streptococcus gordonii* V288 (Challis), GP230, GP232, or GP251.
-
38. The method of claim 33 wherein the host is deficient in extracellular protease production.
- 20 39. The method of claim 33 wherein the gram-positive cell wall surface protein is derived from the genus *Aerococcus*, *Bifidobacterium*, *Corprococcus*, *Deinobacter*, *Deinococcus*, *Enterococcus*, *Gemella*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Marinococcus*, *Melissococcus*, *Micrococcus*, *Pediococcus*, *Peptococcus*, *Peptostreptococcus*, *Planococcus*, *Ruminococcus*, *Saccharococcus*, *Salinococcus*, *Carcina*, *Staphylococcus*, *Stomatococcus*,

-29-

Streptococcus, *Trichococcus*, or *Vagococcus*.

40. The method of claim 39 wherein the gram-positive cell wall surface protein is the type-6 M protein of *Streptococcus pyogenes*.
41. The method of claim 39 wherein the gram-positive cell wall surface protein comprises:
- an N-terminal signal sequence comprising a complete leader sequence and a number of amino acid residues from the N-terminus of the mature protein;
 - a C-terminal sorting signal comprising the LPXTGX motif; wherein X is any amino acid residue, residing within a proline/glycine rich region;
 - a hydrophobic membrane spanning domain; and
 - a charged tail.
42. The method of claim 41, wherein the plasmid vector additionally comprises a DNA fragment having a recognition sequence for at least one restriction endonuclease (the multiple cloning site, MCS) located between the DNA sequences that encode the amino- and carboxy-terminal sorting signals.
43. The method of claim 41 wherein the expression vector additionally comprises a variably sized protein spacer or linker sequence that physically displaces the fusion protein away from the cell surface.
44. The method of claim 43 wherein the linker sequence is incorporated between the protease recognition sequence and the C-terminal sorting signal domain.
45. The method of claim 33 wherein the plasmid vector additionally comprises extrachromosomal genetic markers that are sufficient to provide a selective advantage to a host expressing said genetic markers.

-30-

46. The method of claim 45 wherein the genetic markers are antibiotic genes, heavy metal resistance genes, and nutritional determinants.
47. The method of claim 45 wherein the genetic markers are capable of being expressed in both *E. coli* and gram-positive hosts.
- 5 48. The method of claim 45 wherein the plasmid vector additionally comprises a polylinker or multiple cloning site containing restriction endonuclease cleavage sites.
- 10 49. The method of claim 39 wherein the plasmid vector comprises sequences coding for the M6 ribosome binding site, the translation initiation codon (ATG), the 42 amino acid leader sequence (-42 to -1), and a portion of the mature N-terminus which is at least 5 amino acids long.
50. The method of claim 49 wherein the portion of the N-terminus coded for comprises residues -42 to +16.
- 15 51. The method of claim 33 wherein the vector is one of the SPEX2 series of vectors.
-
52. The method of claim 33 wherein recombinant proteins are purified using metal chelation chromatography.
53. The method of claim 33 wherein non-essential N-terminal residues can be removed with TEV N1a proteinase treatment.

ADANKAQA	LPETGE	ENPLIGTTVEGGLSLALGAALLAGRRREL	Protein A
PMKETKRQ	LPSTGE	TANPFFTAALTVMATAGVAAVVKRKEEN	M6
TKQKAKFV	LPSTGE	QAGLLTTTVGLVIVAVAGVYFTRRR	WapA
AMTQQKRT	LPSTGE	TANPFFTAATAATVMVSAGMLALKRKEEN	M49
PMAQTKRQ	LPSTGE	ETTNPFFTA.....	IgA-BP
DDAKKAET	LPSTGE	GSNPFFTAALAVMAGAGALAVASKRKED	Protein G
KPQSKKSE	LPETGG	EESTNKGMLFGGLFSILGLALILRRNKKNHKA	Fn-BP
IPNTKLGE	LPSTGS	IGTYLFKAIGSAAMIGAIGIYIVKRRKA	T6
QPSSVQET	LPNTGV	TNNAYMPLLGIIGLVTFSFLLGKAKKD	Pac
QLTSGKGA	LPKTGE	TTERPAFGFLGVIVVILMGVLGLKRKQREE	Wg2

Figure 1

2/8

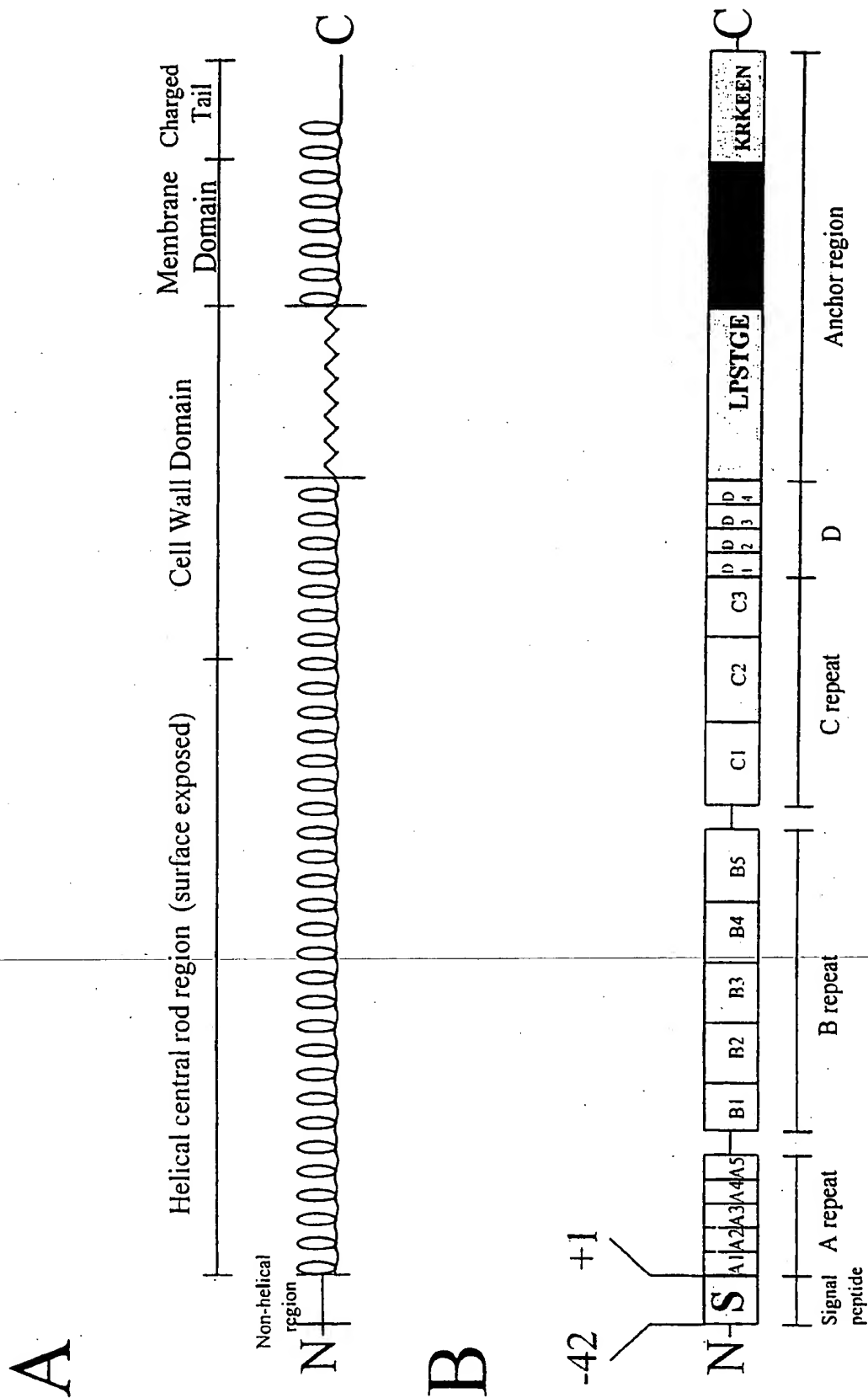


Figure 2

Figure 3

4/8

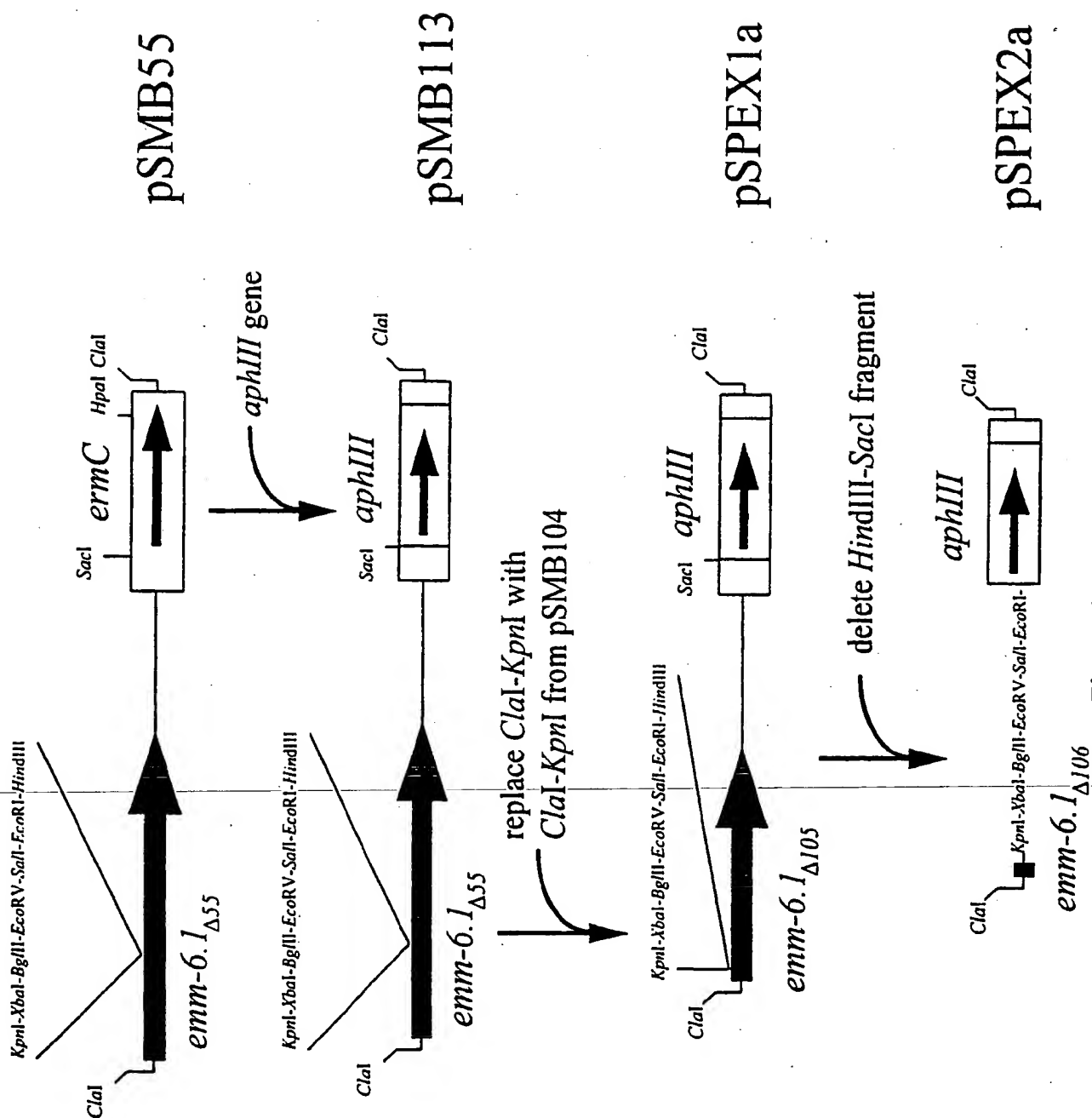


Figure 4

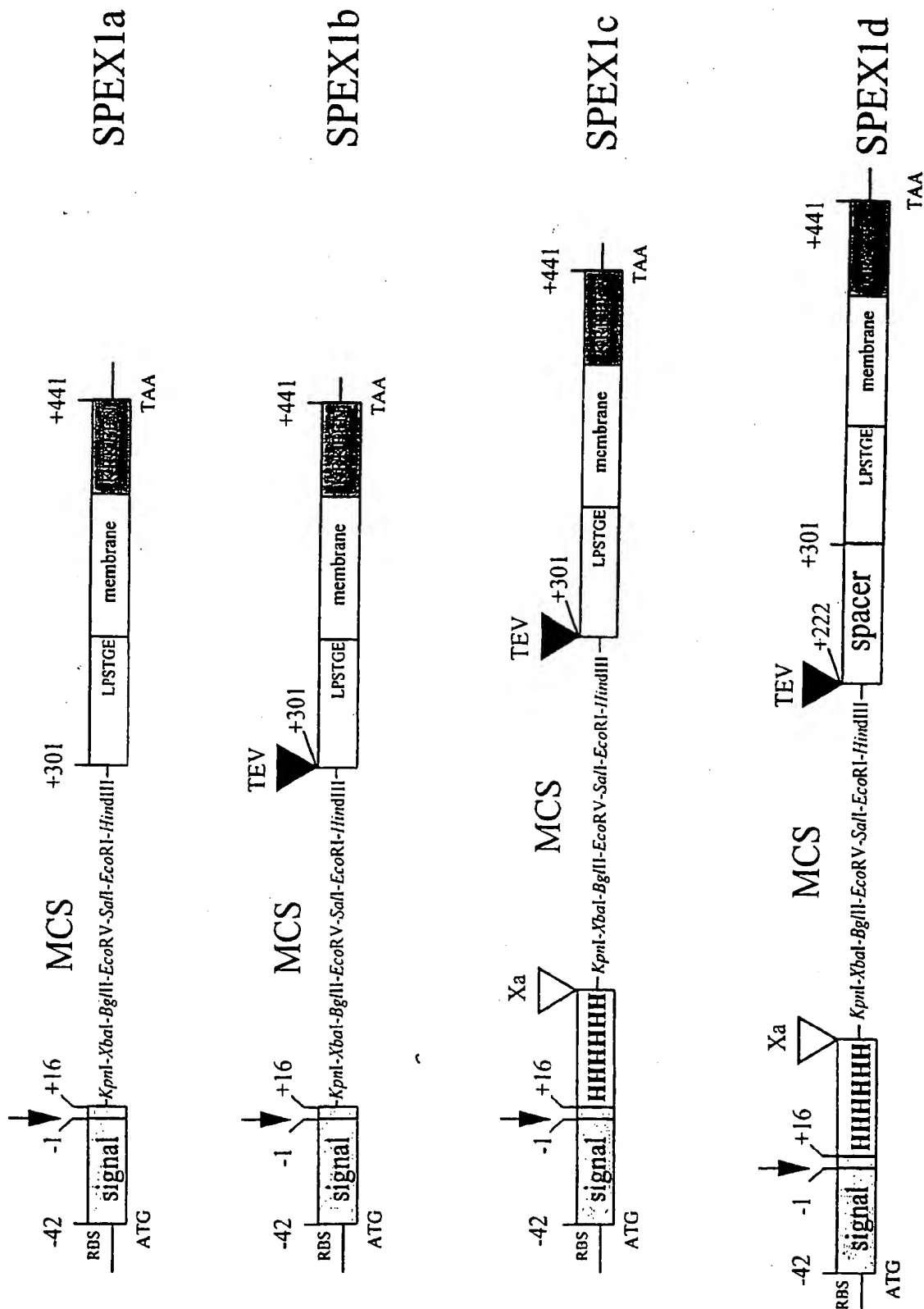


Figure 5



Figure 6

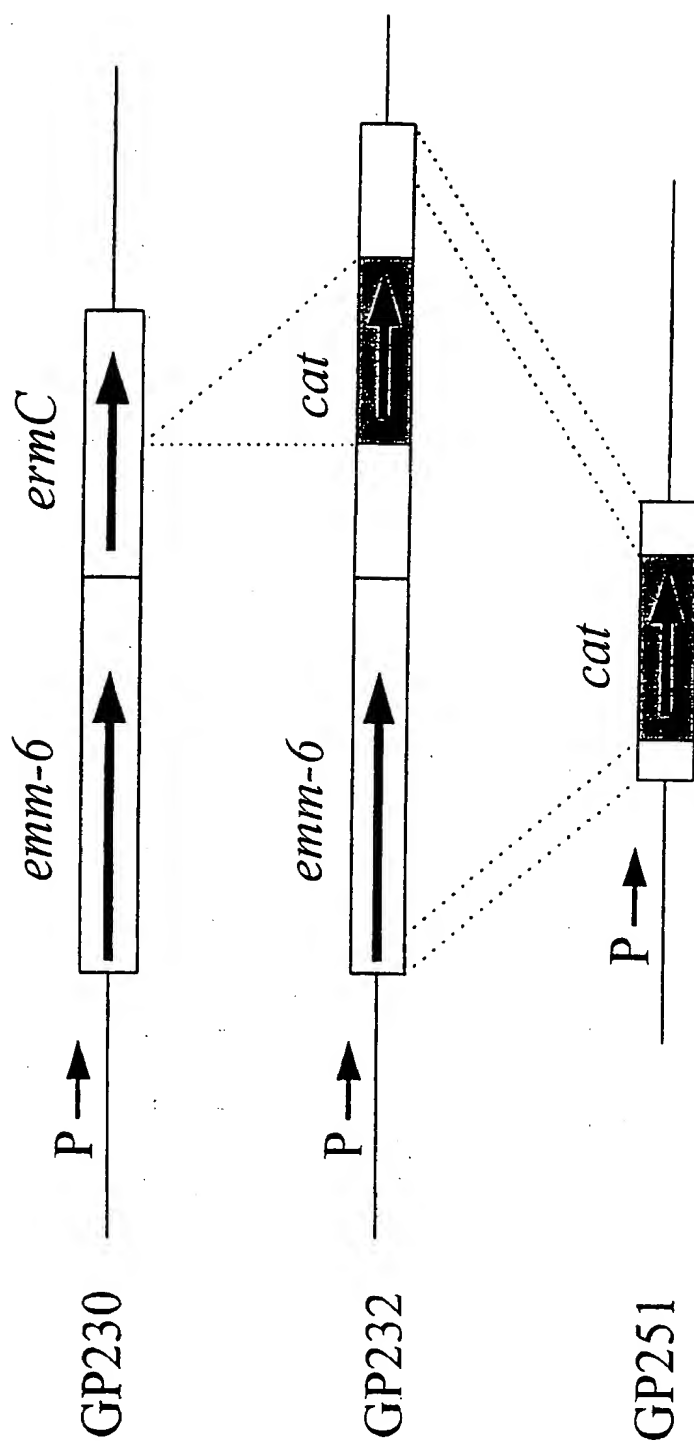


Figure 7

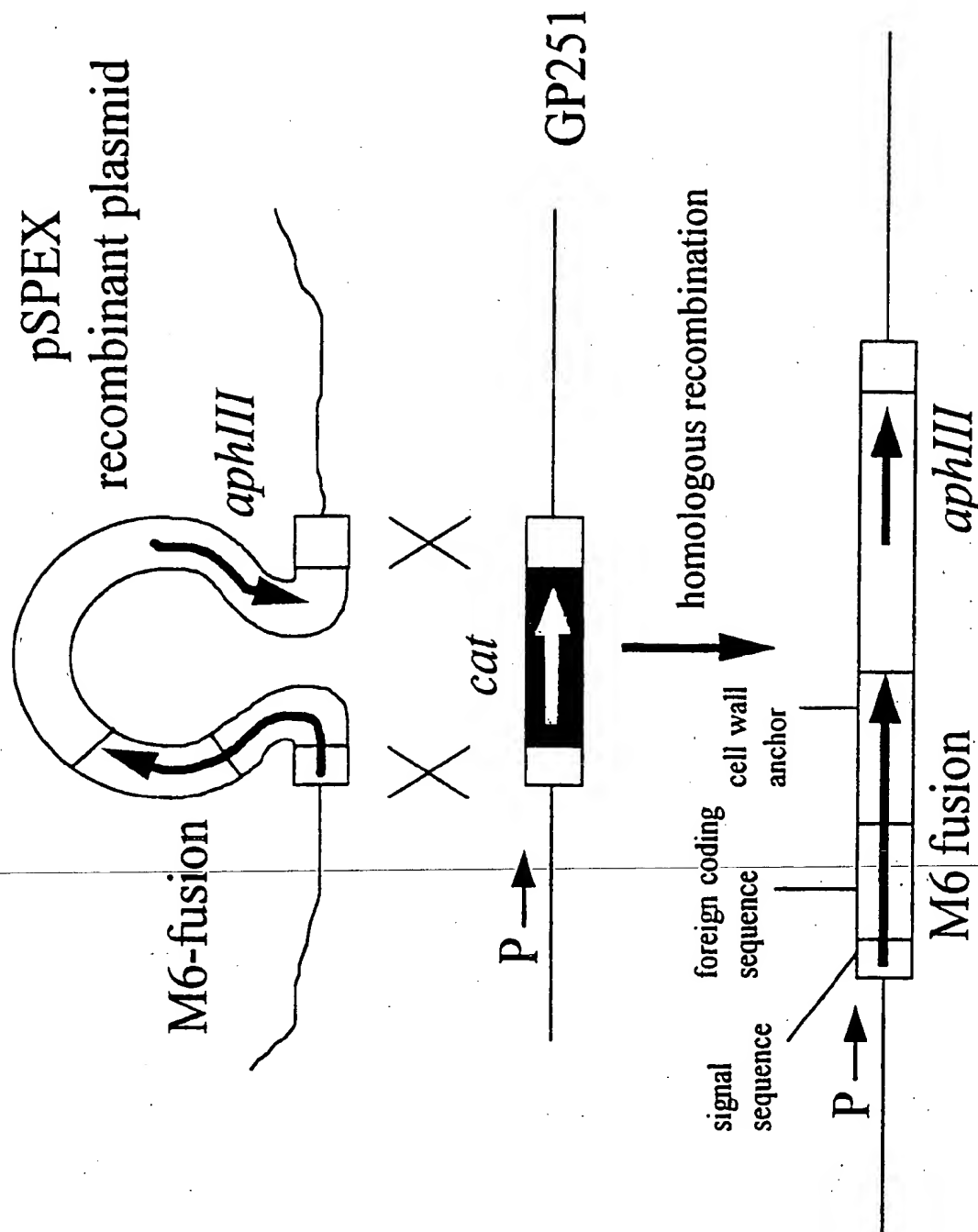


Figure 8

INTERNATIONAL SEARCH REPORT

National Application No

PCT/US 96/09965

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/62 C12N15/31 C12N15/74 C07K14/315

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,93 18163 (UNIV ROCKEFELLER) 16 September 1993 see the whole document ---	1-5, 7-12, 18-24
Y	VACCINE, vol. 12, no. 12, 1994, BUTTERWORTH-HEINEMANN LTD, BOSTON, MA, USA, pages 1071-1077, XP002017663 G. POZZI ET AL.: "Human T-helper cell recognititon of an immunodominant epitope of HIV-1 gp120 expressed on the surface of Streptococcus gordonii" cited in the application see the whole document --- -/--	1-5, 7-12, 18-24

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

5 November 1996

Date of mailing of the international search report

15. 11. 96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+ 31-70) 340-3016

Authorized officer

Hornig, H

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 96/09965

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,92 20805 (PF MEDICAMENT) 26 November 1992 see the whole document ---	1,11,12, 18-21
Y	ANALYTICAL BIOCHEMISTRY, vol. 216, no. 1, 1 February 1994, ACADEMIC PRESS INC., DULUTH,MN,US, pages 413-417, XP002017664 T.D. PARKS ET AL.: "Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase" cited in the application see page 413, right-hand column, line 1 - line 20; figure 1 ---	1-5, 7-12, 18-24
Y	EP,A,0 355 737 (BEHRINGWERKE AG) 28 February 1990 see page 2, line 35 - line 40; claims 6,7 ---	1-5, 7-12, 18-24
Y	WO,A,93 01288 (DEUTSCHES KREBSFORSCH) 21 January 1993 see page 5, line 17 - line 33; claim 7; figure 1 ---	1-5, 7-12, 18-24
A	EMBO J., vol. 12, no. 12, December 1993, OXFORD UNIVERSITY PRESS,GB;; pages 4803-4811, XP002017665 O. SCHNEEWIND ET AL.: "Cell wall sorting signals in the surface proteins of gram-positive bacteria" see the whole document ---	1-53
A	CELL, vol. 70, 24 July 1992, CELL PRESS,CAMBRIDGE,MA,US;; pages 267-281, XP002017666 O. SCHNEEWIND ET AL.: "Sorting of protein A to the Staphylococcal cell wall" cited in the application see the whole document ---	1-53
A	EMBO J., vol. 7, no. 5, 1988, OXFORD UNIVERSITY PRESS,GB;; pages 1281-1287, XP002017667 W.G. DOUGHERTY ET AL.: "Biochemical and mutational analysis of a plant virus polyprotein cleavage site" cited in the application see the whole document -----	1-53

INTERNATIONAL SEARCH REPORT

Information on patent family members

national Application No
PCT/US 96/09965

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9318163	16-09-93	CA-A- 2131995	16-09-93
		EP-A- 0646176	05-04-95
		HU-A- 70306	28-09-95

WO-A-9220805	26-11-92	AU-B- 664841	07-12-95
		AU-A- 1789992	30-12-92
		CA-A- 2103021	01-12-92
		EP-A- 0584167	02-03-94
		JP-T- 7502640	23-03-95
		OA-A- 9866	15-08-94

EP-A-0355737	28-02-90	DE-A- 3828666	01-03-90
		AU-B- 614174	22-08-91
		AU-A- 4016089	01-03-90
		JP-A- 2135095	23-05-90

WO-A-9301288	21-01-93	DE-A- 4122599	04-02-93
		EP-A- 0547201	23-06-93
		JP-T- 6500930	27-01-94

THIS PAGE BLANK (USPTO)



**INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)**

(51) Internationale Patentklassifikation⁶ : C12N 15/30, C07K 14/445, C12N 15/62, A61K 39/015, 31/70, C07H 21/00, C12N 1/21, 5/10, 1/11, 15/67	A3	(11) Internationale Veröffentlichungsnummer: WO 98/14583 (43) Internationales Veröffentlichungsdatum: 9. April 1998 (09.04.98)
(21) Internationales Aktenzeichen: PCT/EP97/05441 (22) Internationales Anmeldedatum: 2. Oktober 1997 (02.10.97) (30) Prioritätsdaten: 196 40 817.2 2. Oktober 1996 (02.10.96) DE (71)(72) Anmelder und Erfinder: BUJARD, Hermann [DE/DE]; Remlerstrasse 9, D-69120 Heidelberg (DE). (72) Erfinder; und (75) Erfinder/Anmelder (nur für US): TOLLE, Ralf [DE/DE]; Friedrich-Naumann-Strasse 8, D-71636 Ludwigsburg (DE). PAN, Weiqing [CN/DE]; Im Buschgewann 71, D-69123 Heidelberg (DE). (74) Anwalt: GRÜNECKER, KINKELDEY, STOCKMAIR & SCHWANHÄUSSER; Maximilianstrasse 58, D-80538 München (DE).		(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Veröffentlicht <i>Mit internationalem Recherchenbericht.</i> <i>Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist. Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i> (88) Veröffentlichungsdatum des internationalen Recherchenberichts: 16. Juli 1998 (16.07.98)
(54) Title: METHOD FOR PRODUCING RECOMBINANTS INTENDED FOR USE IN A COMPLETE MALARIA ANTIGENE GP190/MSP1 (54) Bezeichnung: REKOMBINANTES HERSTELLUNGSVERFAHREN FÜR EIN VOLLSTÄNDIGES MALARIA-ANTIGEN GP190/MSP1 (57) Abstract <p>The present invention relates to a method for producing recombinants intended for use in the complete cell-surface protein gp190/MSP1 from plasmodium, especially plasmodium falciparum, as well as the complete DNA sequence of this protein and the appropriate host organisms suited for expressing said sequence, whereby the protein concerned can be entirely synthesized outside the parasite. Also, the inventive method enables sufficient production of above-mentioned protein and its supply as a vaccine. Finally disclosed is a process for stabilizing genes with high At concentration.</p> (57) Zusammenfassung <p>Die Erfindung betrifft ein rekombinantes Herstellungsverfahren für das komplette gp190/MSP1-Oberflächenprotein von Plasmodium, insbesondere Plasmodium falciparum, sowie die vollständige DNA-Sequenz dieses Proteins und geeignete Wirtsorganismen für die Expression der Sequenz, wodurch das Protein in seiner Gesamtheit erstmals außerhalb des Parasiten synthetisiert werden konnte. Die Erfindung eröffnet erstmals die Möglichkeit, das gp190/MSP1-Oberflächenprotein in ausreichender Menge herzustellen; ferner ist es ein Gegenstand der Erfindung gp190/MSP1 als Impfstoff zur Verfügung zu stellen. Schließlich gibt die Erfindung ein Verfahren zur Stabilisierung von AT-reichen Genen an.</p>		

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidshan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland			TR	Türkei
BG	Bulgarien	HU	Ungarn	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
BY	Belarus	IS	Island	MW	Malawi	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	MX	Mexiko		
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	UZ	Usbekistan
CG	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Jugoslawien
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	NZ	Neuseeland	ZW	Zimbabwe
CM	Kamerun			PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	KZ	Kasachstan	RO	Rumänien		
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		

Rekombinantes Herstellungsverfahren für ein vollständiges Malaria-Antigen gp190/MSP1

Die Erfindung betrifft ein rekombinantes Herstellungsverfahren für das vollständige Malaria-Antigen gp190/MSP1 sowie einzelner natürlich vorkommender Domänen und Teile derselben durch Expression (einer) synthetischer DNA-Sequenzen. Die Erfindung betrifft außerdem die durch das Verfahren hergestellten DNA-Sequenzen und die für die Expression der DNA-Sequenzen verwendeten Wirtsorganismen. Ferner betrifft die Erfindung die Verwendung des vollständigen Malaria-Antigens sowie Teile derselben als Impfstoff zur Immunisierung gegen Malaria.

Schließlich betrifft die vorliegende Erfindung ein Stabilisierungsverfahren für AT-reiche Gene, sowie stabilisierte Gene, die sich durch einen geringeren AT-Gehalt auszeichnen.

Malaria ist weltweit eine der bedeutendsten Infektionskrankheiten. Nach Angaben der WHO waren 1990 in 99 Ländern 40% der Weltbevölkerung einem Malariarisiko ausgesetzt. Ihre Verbreitung nimmt derzeit wieder massiv zu. Dies ist vor allem auf eine intensive Resistenzbildung der Malariaerreger zurückzuführen, die dadurch gefördert wird, daß die zur Therapie eingesetzten Medikamente ebenfalls als Prophylaxe empfohlen und eingenommen werden. Neben der Suche nach neuen, wirksamen Chemotherapeutika werden heute Hoffnungen auch in die Entwicklung von Impfstoffen gesetzt, da Menschen in Malaria-epidemischen Regionen der Welt verschiedene Arten von Immunität zu entwickeln vermögen. Neben einer natürlichen Resistenz gegen Malaria, die sich bei heterozygoten Trägern des Sichelzellgens und bei Personen mit Thalassämie und Glucose-6-Phosphat-Dehydrogenase-Mangel ausbildet, können im Laufe einer Malariainfektion im Menschen Immunitätsmechanismen einsetzen, die sich in einer erhöhten Widerstandsfähigkeit gegenüber den Plasmodien äußern. Dementsprechend ist der Krankheitsverlauf in stark durchseuchten Bevölkerungspopulationen weitaus weniger bedrohlich als bei Personen, die der Infektion weniger häufig oder erstmalig ausgesetzt sind.

Das Hauptproblem bei der Impfstoffentwicklung ist die Identifikation eines Antigens, das schützende Immunität bewirken kann, da kein leicht zugängliches und gut defi-

niertes Tiermodell für die vier den Menschen befallenden Parasiten vorhanden ist. Die Malaria-Erreger gehören der Gruppe der Plasmodien an, wobei die Infektion mit einem der vier Erreger *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* oder *Plasmodium falciparum* durch den Stich von Anopheles-Mücken erfolgt. Von diesem Erreger ist *Plasmodium falciparum* der gefährlichste und der am weitesten verbreitete.

Das Hauptoberflächenprotein von Merozoiten, der invasiven Form der Blutstadien des Malaria-Erregers *Plasmodium falciparum* und anderer Malaria-Erreger wie *P. vivax*, ist ein 190 - 220 kD Glykoprotein. Spät in der Entwicklung des Parasiten wird dieser Vorläufer in kleinere Proteine prozessiert, die jedoch als einheitlicher Komplex aus Merozoiten isoliert werden können. Der Komplex ist mittels eines Glycosylphosphatidyl-Inositol-Ankers mit der Merozoitenmembran verknüpft. Die Sequenzen der gp 190-Proteine verschiedener *P. falciparum*-Stämme fallen in zwei Gruppen, zwischen denen intragene Rekombination häufig ist. Insgesamt besteht das Protein aus mehreren hochkonservierten Regionen, aus einem dimorphen Bereich, welche jeweils einem von zwei Allelen angehören und aus zwei relativ kleinen oligomorphen Blöcken im N-terminalen Bereich (Tanabe, K., Mackay, M., Goman, M. und Scaife, J.G. (1987), Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. J. Mol. Biol. 195, 273-287; Miller, L.H., Roberts, T., Shaha-buddin, M. und McCutchan, T.F. (1993), Analysis of sequence diversity in the *Plasmodium falciparum* merozoite surface protein-1 (MSP-1). Mol. Biochem. Parasitol. 59, 1-14).

Das gp190/MSP1 galt bereits früh als ein möglicher Kandidat für einen Impfstoff. So wurde im Nagermodell nach Immunisierung mit dem analogen Protein aktiver Schutz gegen Infektion mit Nager-Parasiten erhalten. Passiver Schutz ließ sich mit gegen dieses Protein gerichteten Antikörpern erreichen (siehe auch Holder, A. A. und Freeman, R.R. (1981), Immunization against blood-stage rodent malaria using purified parasite antigens, Nature 294, 361-364; Majarian, W.R., Daly, T. M., Weidanz, W.P. und Long, C.A. (1984), Passive immunization against murine malaria with an IgG3 monoclonal antibody, J. Immunol. 132, 3131-3137). Die Daten, die diese Annahme belegen sollen, sind im einzelnen jedoch statistisch nicht signifikant.

Darüber hinaus sind mehrere monoklonale Antikörper, welche in vitro die Invasion von Erythrozyten durch *P. falciparum* inhibieren, sind gegen gp190/MSP1 gerichtet (Pirson, P.J. und Perkins, M.E. (1985), Characterization with monoclonal antibodies of a surface antigen of *Plasmodium falciparum* merozoites. J. Immunol. 134, 1946-1951; Blackman, M.J., Heidrich, H.-G., Donachie, S., McBride, J.S. und Holder A.A. (1990), A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies, J. Exp. Med. 172, 379-382).

Schließlich wurde eine Reihe von Impfstudien mit gp190/MSP1-Material aus *P. falciparum* an Primaten, insbesondere an Aotus und Saimiri-Affen durchgeführt (siehe auch Perrin, L.H., Merkli, B., Loche, M., Chizzolini, C., Smart, J. und Richle, R. (1984), Antimalarial immunity in Saimiri monkeys. Immunization with surface components of asexual blood stages, J. Exp. Med. 160, 441-451; Hall, R., Hyde, J.E., Goman, M., Simmons, D.L., Hope, I.A., Mackay, M. und Scaife, J.G. (1984), Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria, Nature 311, 379-382; Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chan, E.B.T. und Kan, S.-C. (1987), Merozoite surface coat precursor protein completely protects Aotus monkeys against *Plasmodium falciparum* malaria, Proc. Natl. Acad. Sci. USA 84, 3014-3018; Ettlinger, H.M., Caspers, P., Materile, H., Schoenfeld, H.-J., Stueber, D. und Takacs, B. (1991), Ability of recombinant or native proteins to protect monkeys against heterologous challenge with *Plasmodium falciparum*, Inf. Imm. 59, 3498-3503; Holder, A.A., Freeman, R.R. und Nicholls, S.C. (1988), Immunization against *Plasmodium falciparum* with recombinant polypeptides produced in *Escherichia coli*, Parasite Immunol. 10, 607-617; Herrera, S., Herrera, M.A., Perlaza, B.L., Burki, Y., Caspers, P., Döbeli, H., Rotmann, D. und Certa, U. (1990), Immunization of Aotus monkeys with *Plasmodium falciparum* blood-stage recombinant proteins, Proc. Natl. Acad. Sci. USA 87, 4017-4021; Herrera, M.A., Rosero, F., Herrera, S., Caspers, P., Rotmann, D., Sinigaglia, F. und Certa, U. (1992), Protection against malaria in Aotus monkeys immunized with a recombinant blood-stage antigen fused to a universal T-cell epitope; correlation of serum gamma interferon levels with protection, Inf. Imm. 60, 154-158; Patarroyo, M.E., Ro-

mero P., Torres, M.L., Clavijo, P., Moreno, A., Martinez, A., Rodriguez, R., Guzmán, F. und Cabezas, E. (1987), Induction of protective immunity against experimental infection with malaria using synthetic peptides, Nature 328, 629-632). In diesen Impfstudien lassen sich hierbei zwei Ansätze unterscheiden

- Verwendung von aus Parasiten isoliertem Material und
- Einsatz von in heterologen Expressionssystemen gewonnenem Material.

Letzteres bestand in der Regel aus relativ kleinen Teilbereichen des Gesamtproteins. Obwohl die ersten Resultate der in Vorversuchen an Affen durchgeführten Impfungen andeuten, daß gp190/MSP1 einen Schutz vermitteln könnte, haben alle an Primaten durchgeführten Experimente zwei Probleme, welche einen solchen Schluß in Frage stellen:

- (a) sie wurden an zu kleinen Tiergruppen durchgeführt
- (b) sie wurden nicht wiederholt.

Die Resultate und die daraus gezogenen Schlüsse sind daher statistisch nicht abgesichert. Neben dem schwierigen Zugang zu geeigneten Affen liegt das zugrunde liegende Hauptproblem darin, daß es bislang nicht möglich war, gutes Impfmateriel in ausreichender Menge herzustellen.

Andererseits konnten nach der Sequenzierung des gp190-Gens aus dem K1- und dem MAD20-Stamm von *Plasmodium falciparum* konnten überlappende Fragmente in *E. coli* exprimiert werden. Mit diesem Material zeigten epidemiologische Studien in Westafrika, daß in der Gruppe der Adoleszenten eine Korrelation bestand zwischen Antikörpertiter gegen gp190/MSP1-Fragmente einerseits und einem Schutz vor Parasiteninfektion andererseits. Darüber hinaus schien der Titer auch mit der Fähigkeit zu korrelieren, die Parasitämie auch auf niedrigem Niveau zu kontrollieren (Tolle et al. (1993): A prospective study of the association between the human humoral immune response to *Plasmodium falciparum* blood stage antigen gp190 and control of malarial infections. Infect Immun. 61, 40-47). Diese Resultate werden ergänzt durch neue Untersuchungen an Aotusaffen im Rahmen der vorliegenden Erfindung. Hier wurde ein hoher Schutz gegen Infektion mit dem Parasiten dadurch erreicht, daß Proteinpräparationen aus *Plasmodium falciparum*, die überwiegend aus nichtprozessiertem

gp190/MSP1 bestanden, als Impfstoff benutzt worden waren. Die Affen mit dem höchsten Antikörpertiter gegen gp190/MSP1 waren am besten geschützt. Diese Resultate machten letztendlich das gp190 zu einem vielversprechenden Kandidaten für eine Impfstoff gegen *Malaria tropica*.

Von einigen Arbeitsgruppen wurde der C-terminalen Domäne des gp190 (p19 bzw. p42) eine besondere Rolle bei der gp190 vermittelten Immunität zugewiesen (siehe auch Chang, S.P., Case, S.E., Gosnell, W.L., Hashimoto, A., Kramer, K.J., Tam, L. Q., Hashiro, C.Q., Nikaido, C.M., Gibson, H.L., Lee-Ng, C.T., Barr, P.J., Yokota, B.T. und Hui, G.S.N.(1996), A recombinant baculovirus 42-kilodalton C-terminal fragment of *Plasmodium falciparum* merozoite surface protein 1 protects Aotus monkeys against malaria, Inf. Imm. 64, 253-261; Burghaus, P.A., Wellde, B.T., Hall, T., Richards, R.L., Egan, A.F., Riley, E.M., Ripley-Ballou, W. und Holder A.A. (1996), Immunization of Aotus nancymai with recombinant C-terminus of *Plasmodium falciparum* merozoite surface protein-1 in liposomes and alum adjuvant does not induce protection against a challenge infection, Inf. Imm., in press).

Bislang ist es jedoch nicht möglich, auf rationaler Basis andere Teile des gp190 als nicht relevant für eine schützende Immunantwort auszuschließen. Es ist daher nach wie vor notwendig, das gesamte Gen bzw. das intakte gp190 für Impfversuche zu verwenden. Trotz mehrfacher Versuche verschiedener Arbeitskreise ist es jedoch noch nicht gelungen, das gesamte Gen des gp190/MSP1 zu klonieren und zu exprimieren.

Bislang war es auch noch nicht möglich, a priori einen Teil aus der gp190-Sequenz für die schützende Immunantwort als nicht relevant auszuschließen, so daß es nach wie vor notwendig ist, das gesamte Gen bzw. das gesamte Genprodukt für Impfversuche zu verwenden. Es ist jedoch trotz vieler Versuche mehrerer Arbeitskreise noch nicht gelungen, das gesamte Gen des gp190/MSP1 zu klonieren.

Es war daher eine Aufgabe der vorliegenden Erfindung, Impfmateriel in Form von vollständigem gp190/MSP1 in ausreichender Menge zur Verfügung zu stellen. Es war

eine weitere Aufgabe der vorliegenden Erfindung, ein Verfahren anzugeben, mit dem dieses Impfmateriel gewonnen werden kann.

Es war außerdem eine weitere Aufgabe gemäß der vorliegenden Erfindung, eine vollständige DNA-Sequenz von gp190/MSP1 anzugeben, die in einem Wirtsorganismus exprimierbar ist.

Weiterhin war es eine Aufgabe der vorliegenden Erfindung, Wirtsorganismen anzugeben, die das vollständige Gen gp190/MSP1 enthalten.

Schließlich war es auch eine Aufgabe der vorliegenden Erfindung, ein Stabilisierungsverfahren für AT-reiche Gene anzugeben, sowie ein für die Expression geeignetes, stabilisiertes Gen, das sich durch eine Erniedrigung des AT-Gehalts auszeichnet.

Diese Aufgaben werden durch die in den Ansprüchen angegebenen Gegenstände gelöst.

Im folgenden werden einige Begriffe näher erläutert, um klarzustellen, wie sie in diesem Zusammenhang verstanden werden sollen.

"Rekombinantes Herstellungsverfahren" bedeutet, daß ein Protein von einer DNA-Sequenz durch einen geeigneten Wirtsorganismus exprimiert wird, wobei die DNA-Sequenz aus einer Klonierung und Fusion einzelner DNA-Abschnitte entstanden ist.

"Vollständiges gp190/MSP1-Protein" meint hier das gesamte, aus o.g. Plasmodien, insbesondere *Plasmodium falciparum*, isolierbare gp190/MSP1-Oberflächenprotein, das das Hauptoberflächenprotein der Merozoiten des o.g. Erregers darstellt sowie die Proteine mit analoger Funktion aus den anderen Plasmodiumarten, wie *P. vivax*. Der Begriff betrifft somit jeweils das Hauptoberflächenprotein der Merozoiten der 4 o.g. für den Menschen gefährlichen Malariaerreger. "Vollständiges gp190/MSP1-Gen" ist das für dieses Protein kodierende Gen. "Vollständig" bedeutet in diesem Zusammenhang, daß die gesamte Aminosäuresequenz des nativen Proteins vorhanden ist, bzw. daß die Gensequenz für die gesamte Aminosäuresequenz des nativen Proteins kodiert.

Mit umfaßt sind jedoch auch mutierte und/oder verkürzte Formen von gp190/MSP1, sofern sie das gleiche Immunisierungspotential (Impfschutz) wie das vollständige gp190/MSP1 aufweisen. Schließlich umfaßt der Begriff auch Varianten von gp190/MSP1 die sich dadurch auszeichnen, daß sie Proteinabschnitte verschiedener Allele in einem Proteinmolekül enthalten.

"FCB-1" ist ein Stamm von *P. falciparum*, wie beschrieben bei Heidrich, H.-G., Miettinen-Baumann, A., Eckerskorn, C. und Lottspeich, F. (1989) The N-terminal amino acid sequences of the Plasmodium falciparum (FCB1) merozoite surface antigens of 42 and 36 kilodalton, both derived from the 185-195-kilodalton precursor. Mol. Biochem. Parasitol. 34, 147-154.

"Ankersignal" meint hier eine Proteinstruktur, für die eine DNA-Sequenz am 3'- oder 5'-Ende des erfindungsgemäßen Gens kodiert. Ankersignale sind Strukturen, die einem Polypeptid die Verankerung an anderen Strukturen, wie z.B. Membranen ermöglichen.

"Signalpeptid" bedeutet hier eine Proteinstruktur, für die eine DNA-Sequenz am N-terminalen Ende des erfindungsgemäßen Gens kodiert. Signalpeptide sind Strukturen, die dem Polypeptid u.a. ein Einschleusen in Membranen ermöglichen.

"AT-Gehalt" meint im Zusammenhang der vorliegenden Erfindung die prozentuale Menge von Adenin/Thymin-Basenpaaren im Verhältnis zu Guanin/Cytosin-Basenpaaren.

"Klonierung" soll hier alle im Stand der Technik bekannten Klonierungsmethoden umfassen, die hier zum Einsatz kommen könnten, die jedoch nicht alle im einzelnen beschrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

"Expression in einem geeigneten Expressionssystem" soll hier alle im Stand der Technik bekannten Expressionsmethoden in bekannten Expressionssystemen umfassen, die hier zum Einsatz kommen könnten, die jedoch nicht alle im einzelnen be-

schrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

Es ist eine erste Aufgabe gemäß der vorliegenden Erfindung, ein Verfahren anzugeben, durch das das gp190/MSP1-Protein und das Gen hierfür in ausreichender Menge ohne übermäßige Kosten produziert werden kann.

Diese Aufgabe wird durch das in Anspruch 1 beschriebene rekombinante Herstellungsverfahren gelöst, durch das ein voll-ständiges gp190/MSP-1-Gen und das davon kodierte Protein in ausreichender Menge erhältlich ist.

Durch dieses Verfahren wurde es erstmals möglich, das Protein in seiner Gesamtheit außerhalb des Parasiten zu synthetisieren. Das so synthetisierte Protein ist, wie die Analyse mit konformationelle Epitope erkennenden monoklonalen Antikörpern zeigt, zumindest über weite Bereiche in natürlich gefalteter Form herstellbar. Durch das rekombinante Herstellungsverfahren konnten jeweils mehrere Milligramm intaktes gp190/MSP1 aus den Wirtsorganismen gewonnen werden, eine Menge, die aus technischen und aus Kostengründen aus Parasiten praktisch nicht gewonnen werden kann. Die jetzt mögliche Produktion des Proteins in beliebigen Mengen eröffnet neue Perspektiven für seinen Einsatz als experimentellen Impfstoff gegen Malaria. Darüber hinaus ist der Weg frei für die Entwicklung von Lebendimpfstoffen sowie für Vakzine auf Nukleinsäure-Basis.

Vorzugsweise liegt der Synthese der für das Protein gp190/MSP1 kodierenden Gen-Sequenz die Sequenz des *P. falciparum* FCB-1 Stammes zugrunde. *P. falciparum* ist der Erreger der Malaria tropica und damit der gefährlichste unter den Malaria-Arten. Das zugrunde liegende Gen ist ein Vertreter des "K1-Allels", wobei K1 für einen bestimmten *P. falciparum*-Stamm steht. Seine kodierende Sequenz erstreckt sich über 4917 Basenpaare und schließt eine Signalsequenz am N-terminalen Ende sowie eine Anker-Sequenz am C-terminalen Ende ein.

Weiterhin ist das rekombinante Herstellungsverfahren gemäß der Erfindung vorzugsweise dadurch gekennzeichnet, daß der AT-Gehalt der dem Protein zugrunde

liegenden DNA-Sequenz gegenüber der Wildtyp-Sequenz erniedrigt ist, von 74% im ursprünglichen Gen auf vorzugsweise ca. 55%, indem bspw. unter Erhalt der Aminosäuresequenz des FCB-1-Proteins eine DNA-Sequenz mit den im menschlichen Genom üblichen Codon-Häufigkeiten hergestellt wird. Andere Codonhäufigkeiten, welche den AT-Gehalt erniedrigen, sind ebenfalls denkbar.

Vorzugsweise kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz einschließlich Signalpeptid und GPI-Ankersignalpeptid, im weiteren als gp190^S bezeichnet.

In einer weiteren bevorzugten Ausführungsform kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des GPI-Ankersignals. Diese Ausführungsform wird im weiteren als gp190^{S1} bezeichnet.

In noch einer weiteren bevorzugten Ausführung kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des GPI-Ankersignals und des Signalpeptids. Diese Ausführungsform wird im weiteren als gp190^{S2} bezeichnet.

In einer weiteren bevorzugten Ausführungsform kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz und eine Transmembranankersequenz.

In einer besonders bevorzugten Ausführungsform umfaßt das rekombinante Herstellungsverfahren folgende Schritte.

Zunächst den Entwurf der zu synthetisierenden DNA-Sequenz auf der Basis des Gens aus *P. falciparum* FCB-1, wobei eine DNA-Sequenz mit den z.B. im menschlichen Genom üblichen Codon-Häufigkeiten unter Beibehalten der Aminosäuresequenz des FCB-1-Proteins hergestellt wird.

Hierdurch sollte der AT-Gehalt des Gens reduziert werden, vorzugsweise auf 55%.

Im weiteren Verfahren wird die entworfene Sequenz bspw. in 5 überlappende Regionen eingeteilt, welche jeweils Domänen der natürlichen Prozessierungsprodukte des gp190/MSP1-Proteins aus FCB-1 entsprechen: p83, p31, p36, p30 und p19.

Es werden Desoxyoligonukleotide synthetisiert, die jeweils mindestens die gesamte Länge einer Region abdecken.

Besonders bevorzugt werden die Desoxyoligonukleotide so synthetisiert, daß ihre Sequenz abwechselnd dem "oberen" (5' -3') bzw. dem "unteren" (3' - 5') DNA-Strang entspricht. Die Länge dieser Oligonukleotide ist vorzugsweise durchschnittlich 120 Nukleotide und sie überlappt die benachbarten Sequenzen jeweils um ca. 20 Basen.

In einer möglichen Ausführungsform werden DNA-Sequenzen von etwa doppelter Länge wie die jeweiligen Ausgangsprodukte durch asymmetrische PCR hergestellt und zwar so, daß die überschüssigen DNA-Sequenzen, die benachbart sind, jeweils den gegenüberliegenden Strang repräsentieren. Dies führt in einem zweiten

PCR-Amplifikationszyklus zu einem Zweitprodukt, das der Länge von vier ursprünglich eingesetzten Oligonukleotiden (abzüglich der überlappenden Region) entspricht. Die Überführung dieser Produkte in ein überwiegend aus Einzelstrang-DNA bestehendes Präparat durch asymmetrische PCR mit den endständigen Oligonukleotiden erlaubt in einem weiteren Amplifikationsschritt die Herstellung eines 800 bp langen doppelsträngigen DNA-Fragments in nur 25 PCR-Zyklen.

Auf diese Weise werden direkt die kodierenden Regionen für p19, p30, p36 und p31 synthetisiert und in *E. coli* molekular kloniert. Klone mit fehlerfreier Sequenz wurden entweder direkt oder durch Zusammensetzen fehlerfreier Teilsequenzen erhalten. Die Region, welche das p83 kodiert, wurde durch Fusion aus zwei etwa 1200 bp umfassenden Sequenzen erhalten.

Im weiteren Verlauf des Verfahrens wurden die einzelnen Sequenzen kloniert. Als Expressionsvektoren bieten sich vorzugsweise die Plasmide pDS56, RBSII ("Hochuli,

E., Bannwarth, W. Döbeli, H. Gentz, R., and Stüber, D. (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Biotechn. 6*, 1321-1325"), pBi-5 ("Baron, U., Freundlieb, S., Gossen, M. and Bujard, H. (1995) Corregulation of two gene activities by tetracycline via a bidirectional promoter. *Nucl. Acids Res. 23*, 3605-3606") und ppTMCS an. Es sind jedoch auch andere Expressionsvektoren denkbar.

Bevorzugte Wirtsorganismen für die Expression sind *E. coli*, besonders bevorzugt der Stamm DH5alphaZ1 (R. Rutz, Dissertation 1996, Universität Heidelberg), HeLa-Zellen, CHO-Zellen, *Toxoplasma gondii* (Pfefferkorn, E.R. and Pfefferkorn, C.C. 1976, *Toxoplasma gondii*: Isolation and preliminary characterization of temperature - sensitive mutants. *Exp. Parasitol. 39*, 365-376) oder *Leishmania*. Weitere Wirtssysteme könnten z.B. Hefen, Baculoviren oder Adenoviren sein, wobei der Gegenstand der Erfindung nicht auf die genannten Wirtssysteme beschränkt sein sollte.

Es war eine weitere Aufgabe der vorliegenden Erfindung, eine vollständige, zur Expression geeignete DNA-Sequenz des gp190/MSP1-Oberflächenproteins von *P. falciparum* anzugeben.

Diese Aufgabe wird durch die in Anspruch 17 genannte Erfindung gelöst, wobei die Sequenz durch das oben beschriebene rekombinante Herstellungsverfahren erhältlich ist.

In einer bevorzugten Ausführungsform der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz.

In einer anderen bevorzugten Ausführungsform der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz mit Ausnahme des Ankersignals.

In einer weiteren bevorzugten Ausführungsform gemäß der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz mit Ausnahme des Ankersignals und des Signalpeptids. Diese Ausführungs-

form von gp190/MSP1 kann dadurch gekennzeichnet sein, daß sie am N-Terminus 11 zusätzliche Aminosäuren, davon 6 Histidine, enthält.

Besonders bevorzugt enthält die zur Expression geeignete DNA-Sequenz keine erkennbaren "splice-donor" und "splice-acceptor"-Signale, und sie ist vorzugsweise dadurch gekennzeichnet, daß sie keine größeren GC-reichen Sequenzen enthält, die stabile Haarnadelstrukturen auf RNA-Ebene bewirken könnten.

Vorzugsweise sollten Erkennungssignale für Restriktionsenzyme, welche Sequenzen von sechs und mehr Basenpaaren erkennen, vermieden werden.

In einer bevorzugten Ausführung werden spezifische, d.h. nur einmal im Gen vorkommende Schnittstellen für Restriktionsendonukleasen in Regionen eingeführt, welche die nach der Prozessierung des Proteins entstehenden Domänen trennen.

Besonders bevorzugt sollten an beiden Enden des Gens Sequenzen für Restriktionsendonukleasen vorhanden sein, die im Gen nicht vorkommen.

Weiterhin werden durch die Erfindung Wirtsorganismen zur Verfügung gestellt, die die vollständige Sequenz des gp190/MSP-1 Oberflächenproteins enthalten.

Solche Wirtsorganismen sind vorzugsweise *E. coli*, besonders bevorzugt der Stamm DH5alphaZ1, HeLa-Zellen, CHO-Zellen, *Toxoplasma gondii* oder *Leishmania*. Die HeLa-Zellen und CHO-Zellen sollten vorzugsweise konstitutiv tTA synthetisieren.

Schließlich stellt die vorliegende Erfindung eine Möglichkeit zur Verfügung, ein nach dem rekombinanten Herstellungsverfahren erzeugtes gp190/MSP1-Oberflächenprotein oder Teile desselben zur aktiven Immunisierung gegen Malaria zu verwenden.

Das hier vorgestellte Syntheschema erlaubt auch das zweite Allel des gp190/MSP1-Gens herzustellen. Damit wird dem Dimorphismus des Proteins Rechnung getragen. Die Hauptvariabilität des Proteins beruht jedoch auf den Sequenzen von zwei relativ kurzen Blöcken (Block II und IV, Ref. 1), die oligomorph sind. Die vor-

liegenden Sequenzdaten ermöglichen es, daß mit 6-8 Sequenzkombinationen dieser Blöcke über 95% aller bekannten gp190/MSP1-Sequenzen abgedeckt werden können. Die Synthese dieser Sequenzvarianten läßt sich problemlos anhand der hier vorgestellten Strategien verwirklichen, so daß Varianten sowohl in das K1 als auch in das MAD20-Allel eingebaut werden können. Impfstoffe aus den hierdurch entstehenden Sequenzfamilien können ggf. gegen ein breites Spektrum von Parasiten mit gp190/MSP1-Varianten Schutz verleihen.

Die Herstellung von verschiedenen Impfstofftypen ist möglich:

- Auf der Ebene von Proteinpräparaten, wobei jeweils Mischungen der zwei Familien (K1-Typ, MAD20-Typ mit verschiedenen Varianten der Blöcke II und IV) zur Anwendung kommen können. Verschiedene Träger bzw. Adjuvans-Materialien können zum Einsatz kommen: Aluminiumoxid, Liposomen, Iscoms QSz1 etc.
- Auf der Ebene der Lebendimpfstoffe: (a) virale Träger, insbesondere Vakzinia und Adenoviren; (b) Parasiten als Träger, insbesondere avirulente Formen von Leishmania und Toxoplasma; (c) bakterielle Träger, z.B. Salmonella
- Auf der Ebene der Nukleinsäuren, wobei bspw. für die Gentherapie geeignete Vektoren als Vehikel zum Einbringen der Gene in den Wirt verwendet werden; weiterhin ist das Einbringen von Ribonukleinsäuren, welche das gewünschte Protein kodieren, denkbar.

Eine weitere Möglichkeit der Impfung besteht in der Verwendung eines gemäß dem erfindungsgemäßen rekombinanten Herstellungsverfahren erzeugten gp190/MSP1-Proteins zur Herstellung von monoklonalen Antikörpern, die dann ihrerseits zur passiven Immunisierung gegen Malaria verwendet werden.

Ebenso wird eine Verwendung der gemäß dem rekombinanten Herstellungsverfahren in einem Zwischenschritt entstandenen, dem Protein zugrunde liegenden DNA-Sequenz zur Herstellung einer Vakzine auf Nukleinsäurebasis ermöglicht.

Schließlich betrifft die Erfindung auch ein Verfahren zur Stabilisierung von Gen-Sequenzen, insbesondere von Sequenzen, die keine ausreichende Stabilität in Expressionssystemen zeigen.

Diese Stabilisierung wird erfindungsgemäß dadurch erreicht, daß der AT-Gehalt der Sequenz verringert wird.

Weiterhin wird durch die Erfindung ein stabilisiertes Gen zur Verfügung gestellt, das sich dadurch auszeichnet, daß es einen geringeren AT-Gehalt aufweist. Ein Beispiel für ein solches, stabilisiertes Gen ist das Gen für das gp190/MSP1-Oberflächenprotein gemäß der vorliegenden Erfindung.

Im folgenden soll die Erfindung anhand der Abbildungen und Tabellen sowie einiger Beispiele in einzelnen Ausführungsformen beschrieben werden.

Dabei zeigt:

Abb. 1: Schematische Darstellung des gp190/MSP1 Vorläuferproteins aus *P. falciparum* (FCB-1).

Abb. 2: Zwei mit nativem gp190/MSP1 aus *P. falciparum* (FCB-1) an Aotus-Affen durchgeführte Impfversuche.

Abb. 2A: mit 3 x 60 Mikrogramm gp190/MSP1

Abb. 2B: mit 3 x 40 Mikrogramm gp190/MSP1

Abb. 3A: Strategie der Synthese des gp190/MSP1-Gens

Abb. 3B: Prinzip der PCR-gestützten Total-Synthese

Abb. 3C: Totalsequenz des gp190^S

Abb. 3D: N- und C-Terminus der gp190^{S1}-Variante

Abb. 4A: Expressionsvektor pDS56 mit gp190^{S2}-Sequenz

Abb. 4B: Gelelektrophorese von gp190^{S2}.

Abb. 5A: Expressionsvektor pBi-5 mit gp190^{S1}-Sequenz

Abb. 5B: Immunfluoreszenz von HeLa-Zellen

Abb. 5C: Elektrophoretische Charakterisierung von aus HeLa-Zellen aufgereinigtem gp190^{S1}

Abb. 6A: Expressionsvektor ppT 190 mit gp190-Sequenz

Abb. 6B: Immunfluoreszenz der Expression von gp190^S in *T. gondii*

Abb. 6C: Polyacrylamid-Gelelektrophorese von gp190 aus *T. gondii*

Bei dem in Abb. 1 schematisch dargestellten gp190/MSP1-Vorläuferprotein aus *P. falciparum* (FCB-1) repräsentieren die dunklen Blöcke Regionen, die in allen Stämmen hochkonserviert vorliegen. Die schraffierten Blöcke zeigen die dimorphen Bereiche, welche im Falle des FCB-1 Isolates dem K1-Allel entstammen. O1 und O2 zeigen die oligomorphen Bereiche. S bezeichnet die Signalpeptidsequenz, welche 19 Aminosäuren umfaßt, GA, die C-terminale Region, welche das Signal für die GPI-Verankerung des Proteins in der Membran enthält. Die Pfeile deuten die Prozessierungsstellen an, durch die Proteine p53, p31, p36, p30 und p19 entstehen. Das gp190-Gen kodiert insgesamt 1639 Aminosäuren.

Die weiteren Abbildungen werden im Zusammenhang mit den folgenden Beispielen ausführlicher erläutert.

BEISPIELE

Beispiel 1: Totalsynthese einer das gp190/MSP1 kodierenden DNA-Sequenz (siehe hierzu Abb. 3)

A. Strategie der Synthese des gp190/MSP1-Gens (gp190^S) (siehe Abb. 3A).

Die Sequenz wurde aufgeteilt in Fragemente, die den Hauptprozessierungsprodukten entsprachen: p83, p31, p36, p30 und p19. In den Übergangsregionen wurden Spalt-

stellen für Restriktionsendonukleasen (Pfeile in Abb. 3) so eingeplant, daß die Aminosäuresequenz nicht verändert wurde. Alle hier aufgeführten Schnittstellen kommen in der Sequenz nur einmal vor.

Die Fragmente wurden überlappend synthetisiert, so daß die Schnittstellen an den jeweiligen Enden die Verknüpfung zu den benachbarten Fragmenten durch Ligierung möglich machten. Alle Einzelfragmente enthielten zusätzlich an ihrem 5'-Ende eine BamHI-Schnittstelle zur Insertion in Expressionsvektoren. Über MluI und ClaI konnte die Gesamtsequenz kloniert werden. Das hier gezeigte Schema führt zunächst zu einer Sequenz, welche den GPI-Anker nicht ausbilden kann, da C-terminal 18 Aminosäuren fehlen. Die Synthese eines entsprechenden Oligonukleotids, sowie eines über die SphI-Schnittstelle sich erstreckenden "Primers" führt nach PCR zu dem Fragment GA, das über SphI und ClaI eingesetzt werden konnte, die resultierende Totalsequenz war gp190^S. Zur Entfernung der das Signalpeptid kodierenden Sequenz wurden "PCR-Primer" hergestellt, über die das Fragment ΔS synthetisiert wurde. Es erlaubt, über eine BamHI und eine HindIII-Schnittstelle den N-Terminus so zu verändern, daß das Protein mit Aminosäure Nr. 20 begann. Die Kernsequenz, welche das gp190/MSP1 ohne Signalsequenz und ohne GPI-Anker-Signal kodiert, wurde mit gp 190^{S2} bezeichnet. Deletion des GPI-Ankersignals allein führte zu gp190^{S1}.

B. Prinzip der PCR-gestützten Totalsynthese (siehe Abb. 3B)

Oligodesoxynukleotide von etwa 120 Nukleotiden Länge wurden abwechselnd vom kodierenden bzw. nichtkodierenden Strang so synthetisiert, daß sie jeweils etwa 20 Basen mit dem benachbarten Fragment überlappten. Das Schema zeigt beispielhaft die Synthese eines ca. 800 bp langen Fragments aus Oligonukleotiden. In der ersten Stufe wurden in 4 Reaktionsgefäßen jeweils 2 Oligonukleotide "asymmetrisch" amplifiziert. Es entstanden 4 etwa 220 bp lange DNA-Populationen, die vorwiegend aus Einzelsträngen bestanden (A, B, C, D). Vereinigung von A und B sowie von C und D und Amplifikation über 5 Zyklen führte zu 2 etwa 400 bp langen doppelsträngigen Produkten. Asymmetrische Amplifikation dieser DNA-Fragmente (Stufe III) ergab Einzelstrangpopulationen, welche nach Vereinigung und Amplifikation (Stufe IV) nach

10 Zyklen das Endprodukt G von ca. 800 bp Länge ergaben. Diese Synthese war ohne Isolierung der Zwischenprodukte und ohne Puffer oder Enzymerneuerung durchführbar, und war nach 3 Stunden beendet. Das Endprodukt wurde elektrophoretisch gereinigt, mit den geeigneten Restriktionsendonukleasen nachgeschnitten und im pBluescript (Stratagene), in dessen Polylinker eine MluI und eine ClaI-Schnittstelle eingesetzt worden waren, in *E. coli* kloniert.

C. Totalsequenz des gp190^S (siehe Abb. 3C)

Nach Fusion aller Teilsequenzen (Abb. 3A) in pBluescript wurde die Sequenz des Gens mit der Dideoxymethode verifiziert. Das Leseraster des gp190^S hatte eine Länge von 4917 bp (+2 Stopcodons) und kodierte eine Aminosäuresequenz, welche der des gp190/MSP1 aus FCB-1 entspricht (1639 Aminosäuren).

D. N- und C-Terminus der gp190^{S1}-Variante (siehe Abb. 3D)

Die N-terminale Sequenz, beginnend mit der BamHI-Schnittstelle, zeigte den Übergang bei Aminosäure 20, von der angenommen wird, daß sie nach Abspaltung des Signalpeptids den N-Terminus definiert. Am C-Terminus war die kodierte Sequenz bei Aminosäure 1621 zu Ende. Den Stopcodons folgte die ClaI-Schnittstelle.

Beispiel 2: Expression des gp190^{S2} in *E. coli*

A. Expressionsvektor (siehe Abb. 4A)

Die gp190^{S2}-Sequenz wurde über die BamHI- und ClaI-Schnittstellen in pDS56RBSII eingesetzt. Dadurch wurden 6 Histidine sowie einige aus dem Vektor stammenden Aminosäuren an den N-Terminus fusioniert; dies ergibt folgende N-terminale Sequenz des Leserasters; Met Arg Gly Ser (His)₆ Gly Ser. Durch den Promotor P_{N25lacO-1} stand die Transkription unter lacR/O/IPTG-Kontrolle.

B. Expression und Aufreinigung von gp190^{S2} (siehe Abb. 4D)

Eine Überführung des Vektors pDS56RBSIIgp190^{S2} in *E. coli* DH5αZ1 und Induktion der Synthese durch IPTG ergab nach elektrophoretischer Auftrennung des Protein-Totalextraktes der Kultur eine deutlich sichtbare Bande in der erwarteten Größe

(Pfeil). Die Aufreinigung des Materials durch IMAC und Affinitätschromatographie (Antikörpersäule mit mAK5.2) führte zu einem homogenen Produkt von etwa 190 kD. In der Abb. bedeuten M= Molekulargewichtsstandards; 1= *E. coli* Proteine vor, 2= nach Induktion mit IPTG für 2 Stunden. 3,4,5 = Fraktionen aus der Elution der mAK-Säule.

Beispiel 3: Tetrazyklin-kontrollierte Expression von gp190^{S1} in HeLa- und CHO-Zellen und Isolation des Produktes (siehe auch Abb. 5) bzw. 6c)

A. Die gp190-Sequenz wurde in den Expressionsvektor pBi-5 über die BamHI/ClaI-Schnittstellen eingesetzt. Damit stand die Transkription des Gens unter Kontrolle eines bidirektionalen "tTA-responsive"-Promotors und konnte über Tc reguliert werden. Der bidirektionale Promotor initiierte gleichzeitig die Transkription des Indikatorgens Luciferase. Damit ließ sich die Regulation der Expression leicht verfolgen (siehe auch Abb. 5A).

B. Immunfluoreszenz von HeLa-Zellen, welche Luciferase und gp190^{S1} Tc-kontrolliert exprimieren.

In HtTA93-9-Zellen, welche die bidirektionale Transkriptionseinheit von (A) enthalten, wurde mit Antikörpern die Produktion von Luciferase (links), gp190^{S1} (Mitte) in Abwesenheit von Tc nachgewiesen. Nach Zugabe von Tc zeigte sich keine nennenswerte Synthese von gp190^{S1}, (wie dargestellt in Abb 5B, rechts).

C. Elektrophoretische Charakterisierung von aus HeLa-Zellen aufgereinigtem gp190^{S1}.

HeLa-Zellklon HtTA93-9 sowie CHO-Zellklon CHO27-29 wurden mit oder ohne Tc kultiviert. Durch Elektrophorese aufgetrennte Zellextrakte wurden mittels "Western blot" mit mAK5.2 analysiert (Abb. 5C); links zeigt eine Analyse der CHO-, rechts der HeLa-Zelllinie. (1) = Kultur ohne, (2) = Kultur mit Tc, (3) = nicht transfizierte HtTA-1-Zelllinie, Molekulargewichtsstandards sind jeweils links angedeutet.

D. Aufreinigung von in HeLa-Zellklon HtTA93-9 synthetisiertem gp190^{S1}

Die präparative Aufzucht der HtTA-93-9-Linie und Induktion der Expression von der gp190^{S1} durch Tc-Entzug erlaubte die Isolierung des Genproduktes über Affinitätschromatographie (mAK5.2-Säule).

Das Coomassie gefärbte Polyacrylamid-Gel (Abb. 6C) zeigte nach Elektrophorese ein Produkt, das aus gp190^{S1} sowie einem weiteren Protein von ca 50 kD bestand. Letzteres war kein Derivat von gp190^{S1}, stammte also aus HeLA-Zellen. Seine gezielte Abtrennung sollte jedoch keine prinzipielle Schwierigkeit darstellen.

Beispiel 4: Expression von gp190^S in Toxoplasma gondii und Aufreinigung des Produktes (siehe hierzu auch Abb. 6).

A. Die gp190^S-Sequenz wurde über MluI/PstI in den Vektor ppT eingesetzt. Damit wurde das Gen unter die Kontrolle des Tubulin-Promotors (P_{tub-1}) von *T. gondii* gebracht. Die 3' nicht-translatierte Region (VTR) stammte von dem Hauptoberflächenprotein von *T. gondii* (SAG-1) ab.

B. Expression von gp190^S in *T. gondii*.

Transfektion von *T. gondii* mit pTT190 führte zur Isolierung von Parasitenlinien, die konstitutiv gp190^S exprimierten. Die Immunfluoreszenz mit mAK5.2 (mittleres Bild) zeigte nicht nur die Expression des Gens, sondern legte auch die Verankerung des Expressionsproduktes an der Oberfläche des Parasiten nahe, da es, wie SAG-1, das gleiche Immunfluoreszenz-Muster erzeugte (rechter Teil der Abb. 6B); links in Abb. 6B ist eine Phasenkontrastaufnahme des mittleren Bildes dargestellt.

C. Isolierung von gp190^S aus *T. gondii*.

Aus präparativen Mengen von *T. gondii* (5×10^9 Parasiten) wurde gp190^S mittels Affinitätschromatographie (mAK5.2-Säule) aufgereinigt. Das hochreine Protein besaß das erwartete Molekulargewicht, wie das Coomassie-gefärbte Polyacrylamid-Gel nach Elektrophorese zeigte (2-3 aus Abb. 6C). Bei Nr. (1) Abb. 6C ist gereinigtes gp190^{S1} aus CHO-Zellen dargestellt mit Molekulargewichtsmarkierung an der linken Seite.

Beispiel 5: Charakterisierung des gp190^S mit monoklonalen Antikörpern.

Die Wechselwirkung von 16 monoklonalen Antikörpern mit gp190^S aus den verschiedenen heterologen Expressionssystemen wurde durch Immunfluoreszenz (IFA) an *P. falciparum* und *T. gondii* bzw. durch "Western blot" an den aufgereinigten Proteinen überprüft. Völlige Übereinstimmung wurde gefunden, wenn die beiden Parasiten verglichen wurden (Zahl der + deutet die relative Intensität der Fluoreszenz an). Im Western blot reagieren 12 mAK's mit gp190^S aus *E. coli* und *T. gondii*. Im Gegensatz dazu binden 3 Antikörper nicht an das aus CHO-Zellen isolierte Material. Antikörper 15 und 16, die Epitope aus dem oligomorphen bzw. dem alternativen Allel (MAD20) erkennen, reagieren nicht. Die Ergebnisse sind in Tab. 1 zusammengefaßt, wobei ND = nicht durchgeführt bedeutet.

Beispiel 6: Expression des gp190^S in heterologen Systemen.

1. Expression in *E. coli*

Das gp190^{S2} wurde in den Expressionsvektor, pDS56, RBSII eingesetzt, wo es unter Kontrolle des P_{N25lacO-1}-Promotors stand, der über das lac Operator/Repressor/IPTG-System regulierbar ist (Abb. 4A). Die Überführung des Plasmids in Repressor-produzierende *E. coli*-Zellen, z.B. *E. coli* DH5 α Z1 erlaubt es, das gp190^{S2} unter IPTG-Kontrolle zu exprimieren. Das Produkt war aus dem Rohextrakt mittels einer Nickel-Chelatsäule über die durch den Vektor eingebrachte N-terminale (His)₆-Sequenz isolierbar. Eine anschließende Affinitätschromatographie an einer Antikörpersäule führte zu einem hochreinen Präparat. Da der verwendete monoklonale Antikörper (mAk5.2) ein konformationelles Epitop im C-terminalen Bereich erkannte, wird durch diese 2-Stufenreinigung auf intaktes Protein voller Länge, mit korrekter Faltung zumindest am C-Terminus, selektioniert (Abb. 4B).

Das Endprodukt besitzt, im Gegensatz zu dem natürlichen Material, am N-Terminus 11 zusätzliche Aminosäuren, davon 6 Histidine. Es enthält keine N-terminale Signal- und auch keine C-terminale Anker-Sequenz. Die *P. falciparum*-spezifische Sequenz beginnt mit Aminosäure 20 und endet mit Aminosäure 1621.

2. Kontrollierte Expression des gp190^{S1} in HeLa- und CHO-Zellkulturen.

Das gp190^{S1} wurde in den Vektor pBi-5 eingesetzt und damit unter Kontrolle eines durch Tetrazyklin (Tc) regulierbaren Promotors gestellt. Das Tc-kontrollierte System wurde aus 2 Gründen gewählt:

- Es gehört zu den Expressionssystemen, mit denen höchste Ausbeuten in Säugerkulturen erreicht werden.
- Nicht-sekretierte Fremdproteine in hoher Konzentration können mit dem Metabolismus der Zellen in negativer Weise interferieren. Die Synthese des gewünschten Produktes wird daher erst nach Aufwachsen der Kultur initiiert.

In dem Konstrukt pBi5-gp190^{S1} wurde ein bidirektionaler Promotor durch Tc-kontrollierten Transkriptionsaktivator (tTA) aktiviert und initiierte die Transkription sowohl des gp190^{S1} als auch des Luziferase-Indikatorgens. In Anwesenheit von Tc ist der Promotor inaktiv. Die Transkriptionseinheit wurde sowohl in HeLa als auch in CHO-Zellen, welche beide konstitutiv tTA synthetisieren (HtTA-1-Linie (Gossen, M. and Bujard, H. (1992), Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA 89, 5547-5551); CHO-tTA-Linie, unveröffentlicht), überführt. Durch Kotransfektion (Ca²⁺-Phosphat-Methode) mit einem Hygromycin-Resistenz-vermittelnden Markergen wurde auf erfolgreiche chromosomale Integration selektioniert. Hygromycin-resistente Klone wurden dann auf Regulierbarkeit der Expression \geq Tc untersucht, indem die Luziferaseaktivität als Indikator genutzt wurde. In gut regulierbaren Klonen (Regulationsfaktor \leq Tc 1000) wurde die gp190 Synthese geprüft. Immunfluoreszenz-Analyse (Abb. 5B) sowie Untersuchungen über "Western blot" (Abb. 5C) erlaubten, von beiden Zelltypen Klone zu identifizieren, welche gp190 unter streng regulierbaren Bedingungen synthetisieren. Von 20 Klonen wurde jeweils der bestregulierbare subkloniert. Die Subklone HtTA93-9 sowie CH027-29 wurden für Kulturen im 10 l Maßstab verwendet. Aus Zellextrakten dieser Kulturen ließ sich intaktes gp190^{S1} mittels Affinitätschromatographie (mAk5.2) isolieren. Das Material ist homogen bis auf eine einzige zelluläre Komponente, die nicht von gp190^{S1} abstammt und die etwa 25% des Präparats ausmacht (Abb. 6C). Sie müßte in einem weiteren Reinigungsschritt entfernt werden.

3. Expression des gp190^S in Toxoplasma gondii.

Toxoplasma gondii gehört wie *P. falciparum* zu den Apicomplexa und hat daher wahrscheinlich ein dem *P. falciparum* ähnliches Protein-Modifikationssystem. *T. gondii* läßt sich mit Fremd-DNA transfizieren, die effizient in das Genom integriert wird, außerdem läßt sich *T. gondii* problemlos in Zellkultur vermehren. Um ein möglichst natives gp190-Produkt zu erhalten, wurde gp190^{S2} so exprimiert, daß das Protein sekretiert und auf der Oberfläche des Parasiten, wie bei *P. falciparum*, über ein GPI-Motiv in der Membran verankert wird. Dazu wurde das gp190^{S2} (Abb. 3A) in das Plasmid ppTMCS (D. Soldati, unveröffentlicht) eingesetzt (Abb. 6A) und damit unter Kontrolle des *T. gondii*-Tubulin Promotors gestellt.

Dieses Expressionskonstrukt wurde in *T. gondii* transfiziert. Selektion mit Chloramphenicol führte zu resistenten Klonen, die gp190 synthetisieren, wie durch Immunfluoreszenz nachgewiesen wurde (Abb. 6B). Die Immunfluoreszenz mit anti-gp190-Antikörpern war nicht unterscheidbar von einer entsprechenden Anfärbung der Parasiten mittels Antikörper gegen SAG1, dem Hauptoberflächenprotein von *T. gondii*. Es ist daher davon auszugehen, daß gp190 an der Oberfläche von *T. gondii* verankert ist. Mehrere *T. gondii*-Klone (Nr. 3.1 bis 3.4) wurden charakterisiert und für die Produktion von gp190 aufbewahrt. Aus in präparativem Maßstab aufgewachsenen *T. gondii*-Kulturen (Klon 3.4) wurde gp190 mittels Affinitätschromatographie (mAK5.2.-Säule) isoliert. Analyse im elektrischen Feld zeigte ein homogenes Produkt mit einer Wanderungsgeschwindigkeit, die auf das intakte Protein schließen läßt (Abb. 6C).

Beispiel 7: Charakterisierung von gp190 Protein aus verschiedenen Expressionssystemen mittels monoklonaler Antikörper.

Ein Satz gp190-spezifischer monoklonaler Antikörper, von denen mehrere konformationelle Epitope erkennen, wurde dazu benutzt, über Immunfluoreszenz die Reaktivität der Antikörper mit *P. falciparum*- bzw. *T. gondii*-Parasiten zu vergleichen. Tabelle 1 zeigt, daß die Reaktivität der 16 Antikörper für beide Parasiten gleich ist. Dies ist ein starker Hinweis darauf, daß in *T. gondii* weitestgehend "natives" gp190 produziert wird. Der Vergleich der Reaktivität der Antikörper mit Protein aus *E. coli*, HeLa- bzw. CHO-Zellen sowie *T. gondii* zeigt ebenfalls, daß die meisten Antikörper mit den 4 Präparaten reagieren. Insbesondere wird das aus *E. coli* isolierte Protein von mehr

Antikörpern erkannt als das Produkt aus Säugerzellen. Dies ist wahrscheinlich eine Konsequenz der Glykosylierung in Säugerzellen.

Beispiel 8: Immunsierung von Aotus lemurinus griseimembra-Affen mit gp190/MSP1 aus *P. falciparum* (FCB-1).

Zwei unabhängige Immunsierungsexperimente (A, B) wurden durchgeführt. Dazu wurde aus jeweils ca. 2×10^{11} Parasiten unter schonenden Bedingungen einmal 1,0 mg (A) und einmal 0,6 mg hochreines gp190/MSP1 isoliert.

Das Protein wurde zusammen mit Freund'schem Adjuvans (FCA) verabreicht. Die Kontrollgruppe erhielt lediglich FCA. Es wurde dreimal im Abstand von 4 Wochen mit gleicher Menge an Protein bzw. mit Adjuvans immunisiert. Zwei Wochen nach der letzten Immunisierung wurden die Tiere mit jeweils 10^5 Parasiten (FVO-Stamm) aus einem Donor-Tier infiziert. Die Parasitämie wurde täglich gemessen. Die Ergebnisse sind in Abb. 2 zusammengefaßt. Dabei bedeutet

T: daß die Tiere mit Resochin behandelt wurden

D: ein verstorbenes Tier

Abb. 2A.: die Individuen der geimpften Gruppe erhielten je 3 x 60 Mikrogramm gp190/MSP1

Abb. 2B: die Individuen der geimpften Gruppe erhielten je 3 x 40 Mikrogramm gp190/MSP1

Während in den Kontrollgruppen nur 1/11 Tiere keine Parasitämie entwickelten, waren es in der geimpften Gruppe 6/10. Die vier Tiere aus der geimpften Gruppe, die eine hohe Parasitämie entwickelten, taten dies - im Vergleich zur Kontrollgruppe - mit einer durchschnittlichen Verzögerung von vier Tagen (Überschreiten der 2%-Grenze der Parasitämie).

Diese Experimente zeigten erstmals einen hochsignifikanten Schutz gegen Infektion mit *P. falciparum* durch gp190/MSP1 im Affenmodell. Das erfindungsgemäße Verfahren erlaubt es somit erstmalig, einen wirksamen Impfstoff gegen die Malaria anzugeben.

Tabelle 1: Wechselwirkung von gp190^S mit monoklonalen Antikörpern

Code	mAb	Art des Epitops	Variabilität	IFA			Western blot		
				P.f. FCB	Toxoplasma	<i>E. coli</i>	Toxo-plasma	CHO	
1	5,2	konformationell	konserviert	++++	++++	+	+	+	
2	12,10	konformationell	konserviert	++++	++++	+	+	+	
3	7,5	konformationell	konserviert	++++	++++	+	+	+	
4	12,8	konformationell	konserviert	++	++	+	+	+	
5	7,3	konformationell	dimorph (K1)	++++	+++	+	+	+	
6	2,2	konformationell	konserviert	++++	++++	+	+	+	
7	7,6	konformationell	dimorph (K1)	++++	++++	+	+	+	
8	9,8	konformationell	konserviert	++++	++	+	+	-	
9	13,2	sequentiell	konserviert	++++	++++	+	+	+	
10	13,1	sequentiell	dimorph (K1)	++++	+++	+	+	-	
11	6,1	sequentiell	dimorph (K1)	++++	++++	+	+	ND	
12	A5Z	nicht bekannt	nicht bekannt	+++	+++	+	+	+	
13	17,2	nicht bekannt	nicht bekannt	++++	+++	ND	ND	ND	
14	15,2	nicht bekannt	nicht bekannt	++++	+++	ND	ND	ND	
15	9,7	konformationell	dimorph (MAD20)	-	-	-	-	-	
16	12,1	sequentiell	oligomorph	-	-	-	-	-	

Patentansprüche

1. Herstellungsverfahren für das vollständige gp190/MSP1-Protein von Plasmodium, insbesondere *Plasmodium falciparum*, **dadurch gekennzeichnet**, daß das vollständige Gen für gp190/MSP1 in einem geeigneten System, vorzugsweise einem Wirtorganismus, exprimiert wird.
2. Rekombinantes Herstellungsverfahren gemäß Anspruch 1, **dadurch gekennzeichnet**, daß der Synthese der dem Protein zugrunde liegenden DNA-Sequenz die DNA-Sequenz des *P. falciparum*-Stammes FCB-1 zugrunde gelegt wird.
3. Rekombinantes Herstellungsverfahren gemäß Anspruch 2, **dadurch gekennzeichnet**, daß der AT-Gehalt der exprimierten DNA-Sequenz gegenüber der natürlich vorkommenden Sequenz reduziert wurde, vorzugsweise von 74% auf 55%.
4. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz einschließlich Signalpeptid und Ankersignal kodiert.
5. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals kodiert.
6. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals und des Signalpeptides kodiert.
7. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 6, **dadurch gekennzeichnet**, daß es folgende Schritte umfaßt:

- (a) Entwurf der zu synthetisierenden DNA-Sequenz aus *P. falciparum* FCB-1, wobei eine DNA-Sequenz mit den im menschlichen Genom üblichen Codon-Häufigkeiten unter Erhalt der Aminosäuresequenz des FCB-1 Proteins hergestellt werden sollte,
 - (b) Einteilung der entworfenen Sequenz in überlappende Regionen, vorzugsweise in Regionen p83, p31, p36, gp30 und gp19,
 - (c) Synthese von Desoxyoligonukleotiden, die jeweils mindestens die gesamte Länge einer Region abdecken,
 - (d) Synthese der kodierenden Regionen für gp19, gp30, p36 und p31 durch PCR und Synthese der kodierenden Region für p83 durch Fusion aus zwei etwa 1200bp umfassenden Sequenzen,
 - (e) einzelne Klonierung der kodierenden Sequenzen
 - (f) Fusion des gesamten Gens und
 - (g) Expression in einem geeigneten Expressionssystem.
8. Rekombinantes Herstellungsverfahren nach Anspruch 7, **dadurch gekennzeichnet**, daß die in Schritt (c) synthetisierten Desoxyoligonukleotide durchschnittlich 120 Nukleotide lang sind und die benachbarten Sequenzen jeweils um ca 20 Basen überlappen.
-
9. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor dPS56, RBSII verwendet wird.
10. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor pBi-5 verwendet wird.
11. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor ppTMCS verwendet wird.

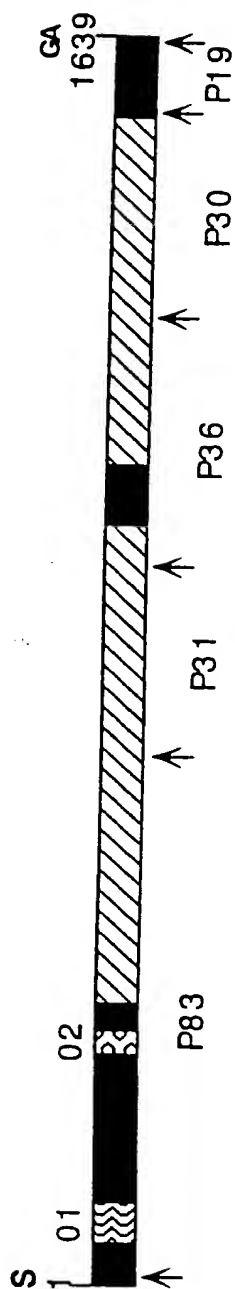
12. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in *E. coli* exprimiert wird.
13. Rekombinantes Herstellungsverfahren nach Anspruch 12, **dadurch gekennzeichnet**, daß der verwendete *E. coli*-Stamm der Repressor-produzierende Stamm *E. coli* DH5alphaZ1 ist.
14. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in HeLa-Zellen exprimiert wird.
15. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in CHO-Zellen exprimiert wird.
16. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in *Toxoplasma gondii* oder *Leishmania* exprimiert wird.
17. Vollständige, zur Expression geeignete DNA-Sequenz des gp190/MSP1-Oberflächenproteins von Plasmodium, insbesondere *P. falciparum*, vorzugsweise erhältlich durch das rekombinante Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 16.
18. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 17, **dadurch gekennzeichnet**, daß sie nicht für das Ankersignal kodiert.
19. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 18, **dadurch gekennzeichnet**, daß sie weder für das Ankersignal noch für das Signalpeptid kodiert.

20. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 19, **dadurch gekennzeichnet**, daß sie eine Sequenz für am N-Terminus vorliegende 11 zusätzliche Aminosäuren, davon 6 Histidine, umfaßt.
 21. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 20, **dadurch gekennzeichnet**, daß die Sequenz keine erkennbaren "splice donor"- und "splice acceptor"-Signale enthält.
 22. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 21, **dadurch gekennzeichnet**, daß die Sequenz keine größeren GC-reichen Sequenzen enthält.
 23. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 22, **dadurch gekennzeichnet**, daß die Sequenz keine Erkennungssignale für Restriktionsenzyme, welche Sequenzen von sechs oder mehr Basenpaaren erkennen, enthält.
 24. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 23, **dadurch gekennzeichnet**, daß die Sequenz für Erkennungssignale bestimmter Restriktionsnukleasen in Regionen, die die nach der Prozessierung des Proteins entstehenden Domänen trennen, einmal vorkommende Schnittstellen für Restriktionsendonukleasen enthält.
-
25. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 24, **dadurch gekennzeichnet**, daß die Sequenz an ihren beiden Enden Schnittstellen für Restriktionsendonukleasen aufweist, die in der übrigen Sequenz und in einem zu verwendenden Vektor nicht vorkommen.
 26. Wirtsorganismus, der die vollständige Nukleinsäuresequenz für das gp190/MSP1-Oberflächenprotein und/oder das vollständige Protein enthält.
 27. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus *E. coli* ist.

28. Wirtsorganismus gemäß Anspruch 27, **dadurch gekennzeichnet**, daß der *E. coli*-Stamm der Repressor-produzierende *E. coli*-Stamm DH5alphaZ1 ist.
29. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus HeLa-Zellen sind.
30. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus CHO-Zellen sind.
31. Wirtsorganismus gemäß Anspruch 29 oder 30, **dadurch gekennzeichnet**, daß die Wirtszellen konstitutiv tTA synthetisieren.
32. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus *Toxoplasma gondii*, *Leishmania*, Baculoviren, Adenoviren oder Hefen vorgesehen sind.
33. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur aktiven Immunisierung gegen Malaria.
34. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur Herstellung von monoklonalen, zur passiven Immunisierung geeigneten Antikörpern.
35. Verwendung einer gemäß der Ansprüche 1 bis 16 hergestellten DNA-Sequenz zur Herstellung einer Vakzine auf Nukleinsäurebasis.

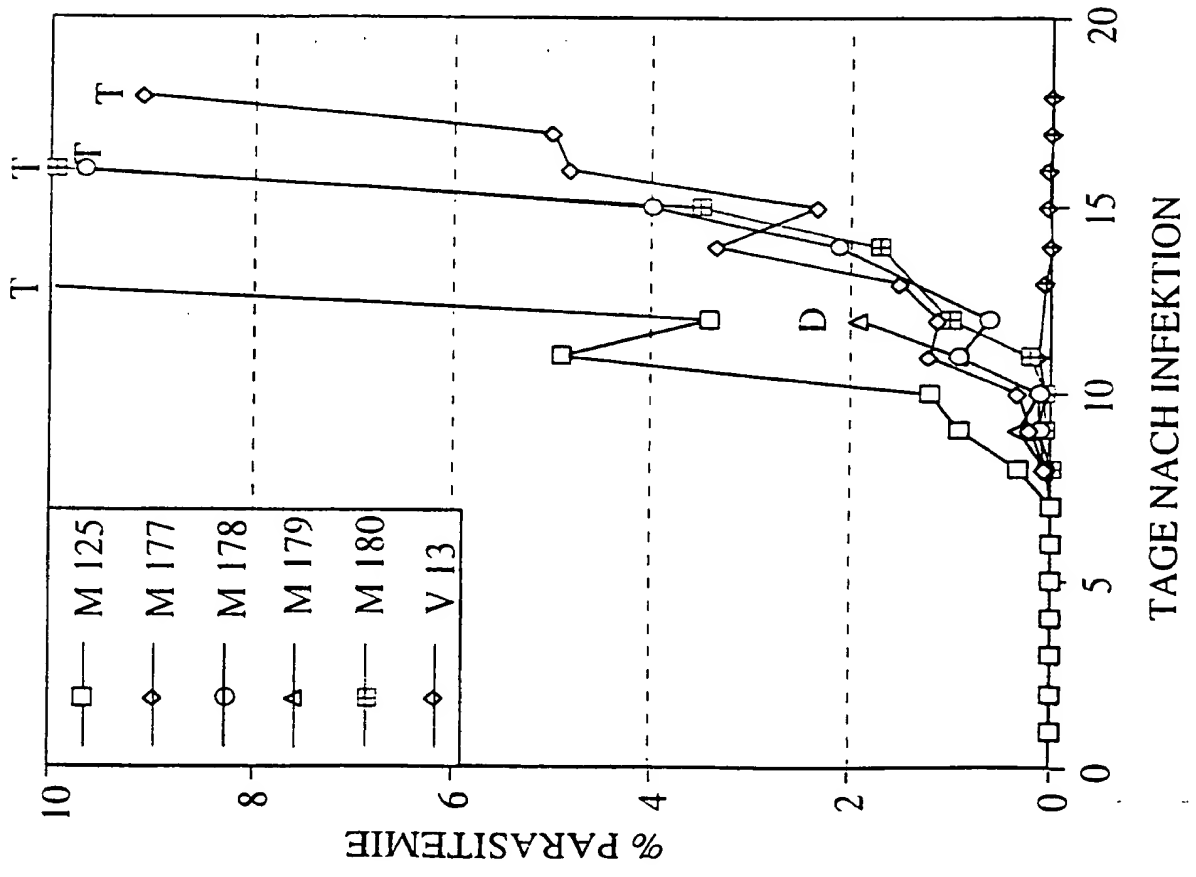
36. Verfahren zur Stabilisierung von Gen-Sequenzen, **dadurch gekennzeichnet**, daß der AT-Gehalt der Sequenz verringert wird.
37. Stabilisiertes Gen, **dadurch gekennzeichnet**, daß es einen geringeren AT-Gehalt aufweist als das nicht stabilisierte Gen.
38. Vektor enthaltend eine DNA-Sequenz nach einem der Ansprüche 17 bis 25 und/oder 37.
39. Wirtszelle enthaltend einen Vektor nach Anspruch 38.
40. Impfstoff enthaltend ein Protein, hergestellt nach einem Verfahren nach einem der Ansprüche 1-16 und/oder eine DNA-Sequenz nach einem der Ansprüche 17-25 und/oder einen Wirt nach einem der Ansprüche 26-32 und/oder einen Vektor nach Anspruch 38.
-
41. Impfstoff nach Anspruch 40, **dadurch gekennzeichnet**, daß er weitere Immunität hervorrufende Produkte aus Plasmodium, insbesondere *P. falciparum*, enthält.

Abb.1

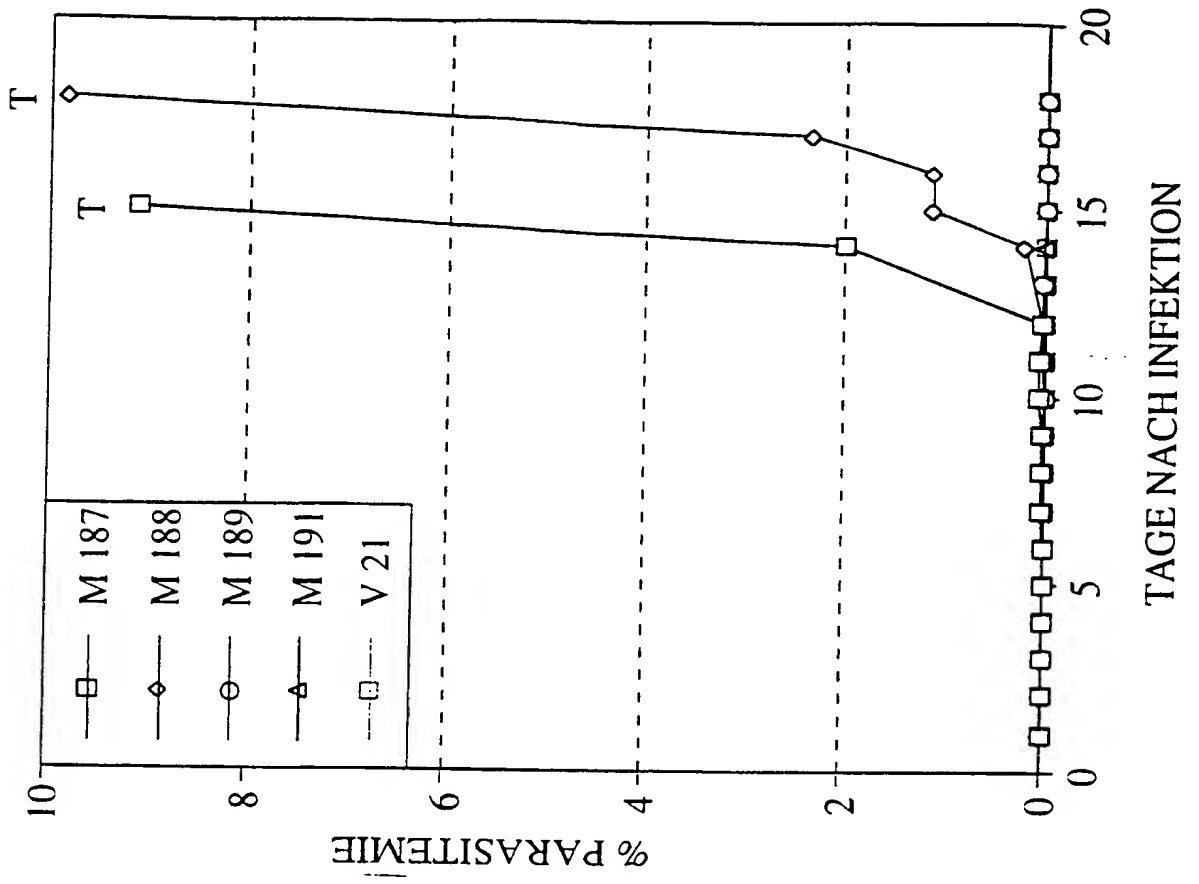


THIS PAGE BLANK (USPTO)

Abb. 2A KONTROLLGRUPPE

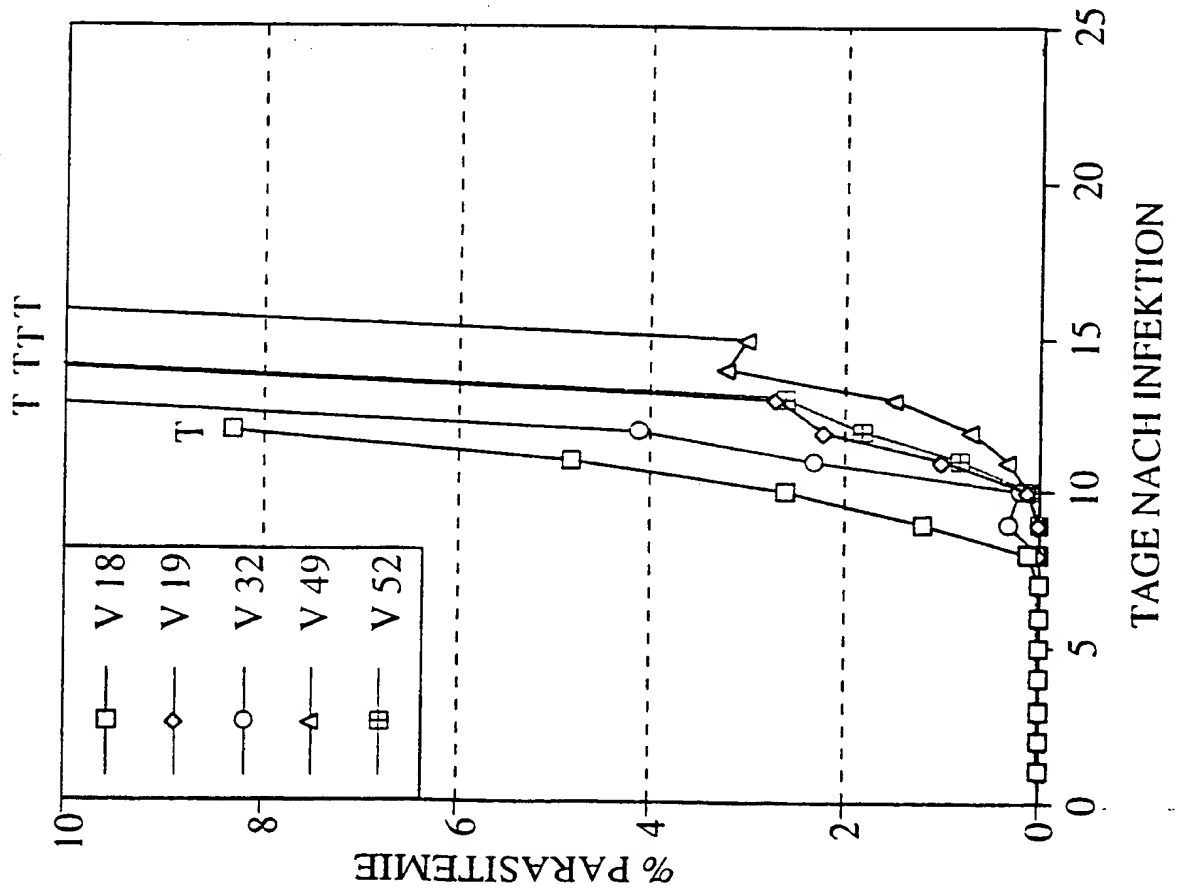


gp190 GRUPPE

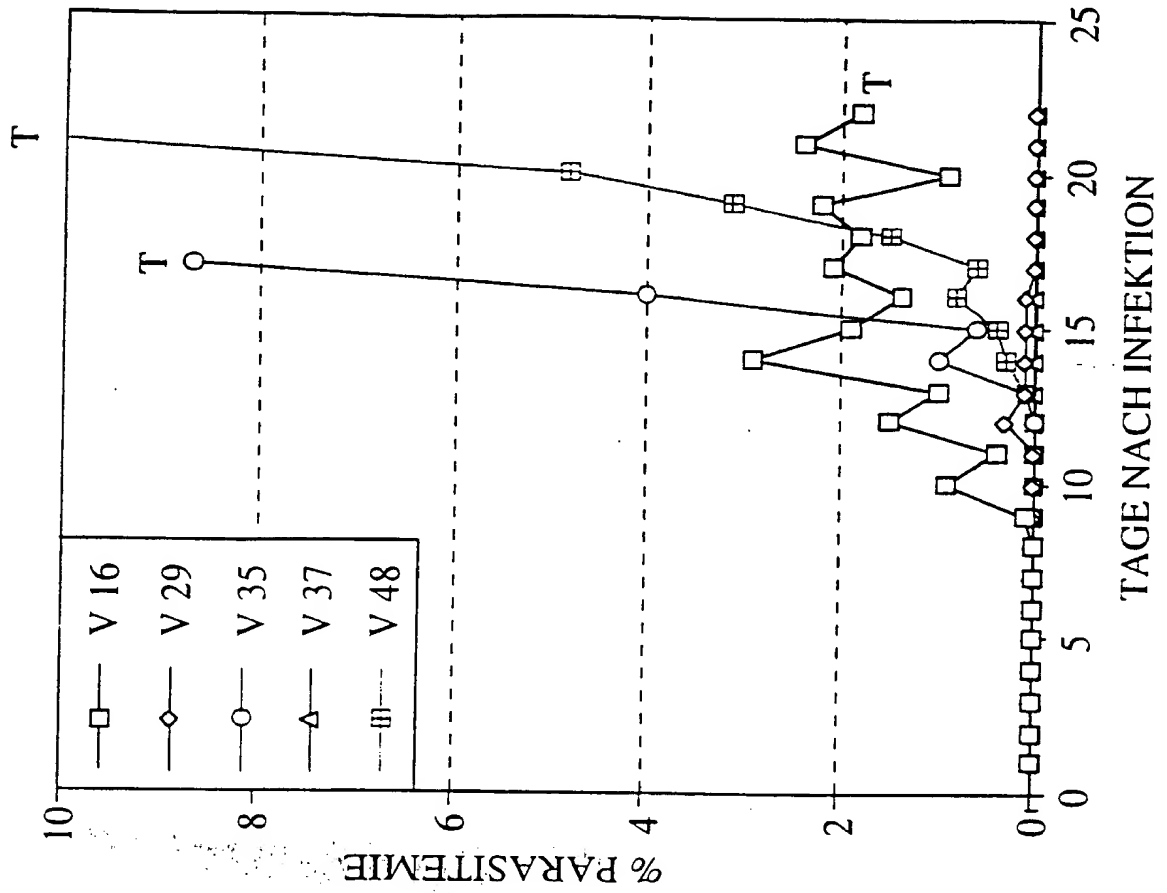


THIS PAGE BLANK (USPTO)

Abb. 2B KONTROLLGRUPPE

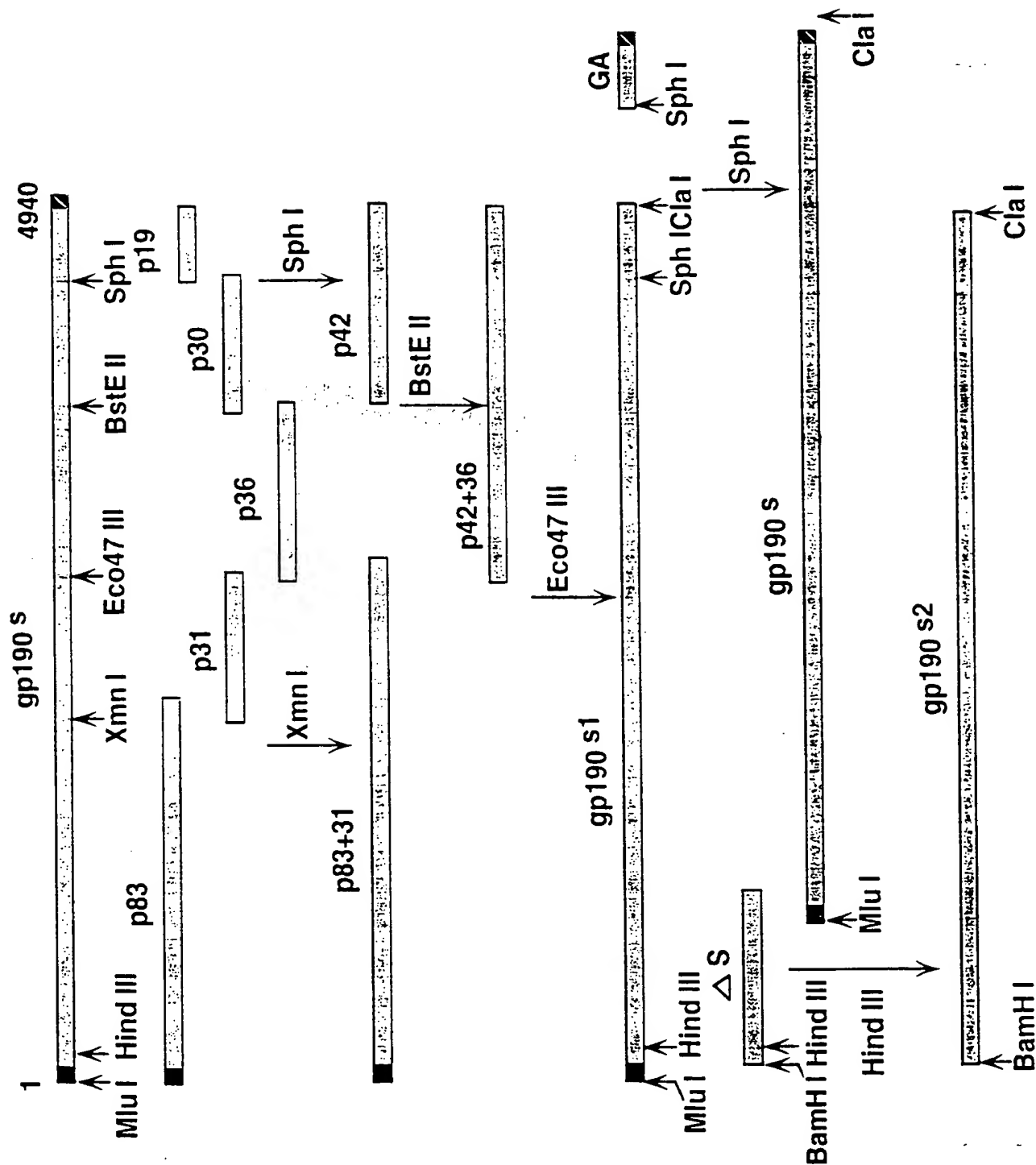


gp190 GRUPPE



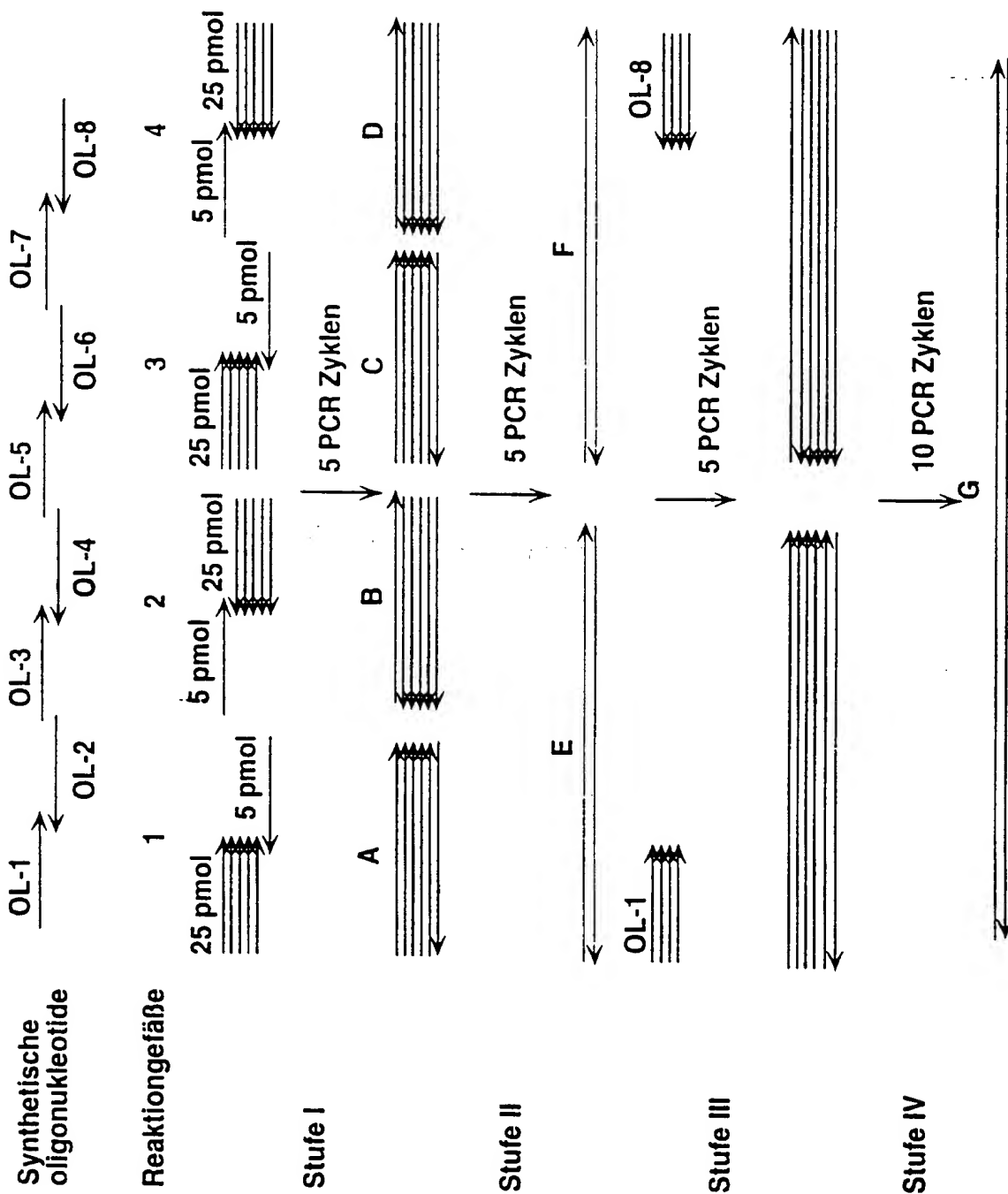
THIS PAGE BLANK (USPTO)

Abb.3A



THIS PAGE BLANK (USPTO)

Abb.3B



THIS PAGE BLANK (USPTO)

Abb. 3C

DNA Sequenz des nativen (gp190ⁿ) und des synthetischen (gp190^s) Gens
für gp190 aus FCB-1

AS	M	K	I	I	F	F	L	C	S	F	L	F	12			
gp190 ⁿ																
			G	A			TT	A				T				
gp190 ^s													45			
	CGCACGCGTATGAA	AAATCATTTTCTTCC	CTCTGTT	CATTCT	CTGTT											
	Mlu I															
AS	F	I	I	N	T	Q	C	V	T	H	E	S	Y	Q	E	27
gp190 ⁿ																
gp190 ^s																90
	TTTATCATCAATA	CTACTCAGTGC	GTGACCCAC	GAAATCCT	ATCAGGAG											
AS	L	V	K	K	L	E	A	L	E	D	A	V	L	T	G	42
gp190 ⁿ																
gp190 ^s																135
	CTGGTTAAGAA	AACTGGAAGCTTT	GGAAGATGCC	CGTCTTAC	CCGA											
AS	Y	S	L	F	Q	K	E	K	M	V	L	N	E	G	T	57
gp190 ⁿ																
gp190 ^s																180
	TACAGCCCTGT	TCCAGAGGAG	AAGATGGT	GCTGAAT	GAAGGACG											

THIS PAGE BLANK (USPTO)

```

AS      S G T A V T T S T P G S K G S      72
gp190n  A A T      T T      A G T A
gp190S  AGTGCACGGCCGTTACAACCAGCACACCCGGTTCTAAAGGTCT 225

AS      V A S G G S G G S V A S G G S      87
gp190n  T TCA T A C A T T A T C A
gp190S  GTGGCTAGCGGTGGCTCCGGTGGTCTGTGGCCTCTGGGGTTCC 270

AS      V A S G G S V A S G G S V A S      102
gp190n  T T A T TCA T T      T TTCA
gp190S  GTCGCCCTCCGGCGCAGCGTGGCATCAGTGGCTCAGTGGCAAGC 315

AS      G G S G N S R R T N P S D N S      117
gp190n  T A T TTCAA C T A T A T T A
gp190S  GCGGTTCCGGGAACAGTCGAAGAACCAATCCATCTGACAACTCT 360

AS      S D S D A K S Y A D L K H R V      132
gp190n  T A T T A T T TT A A A A
gp190S  AGCGATTCCGACGCCAAGTCCCTACGCCGACCTCAAGCACCGAGTG 405

AS      R N Y L L T I K E L K Y P Q L      147
gp190n  C T CT GT A A A C A T T AC C
gp190S  AGAAACTATCTCCTCACTATCAAGGAGCTGAAGTACCCACAGTTG 450

```

THIS PAGE BLANK (USPTO)

8/36

AS	F	D	L	L	T	N	H	M	L	T	L	C	D	N	I	H	162
gp190n	T	TT	A						T	A	TT						
gp190S	TT	G	A	C	C	T	C	A	C	T	A	T	A	T	G	T	495
AS	G	F	K	Y	L	I	D	G	Y	E	E	I	N	E	L		177
gp190n	T				T	A		T	A	T		A	T		T	A	
gp190S	GG	CT	CA	AA	T	A	T	C	T	G	A	T	T	G	A	T	540
AS	L	Y	K	L	N	F	Y	F	D	L	L	R	A	K	L		192
gp190n	T	A	T	A	A	C	T	T	T	T	AT	A	A	T	A		
gp190S	CT	GT	ACA	AGT	TGA	ATT	CT	ACT	TC	G	ACT	TG	CT	A	AGG	CCAA	585
AS	N	D	V	C	A	N	D	Y	C	Q	I	P	F	N	L		207
gp190n	T	A	T	T			T			A	T			C	T		
gp190S	AA	TG	AC	GT	TTG	CG	CC	AA	TG	ACT	AT	TG	T	CA	AA	TT	630
AS	K	I	R	A	N	E	L	D	V	L	K	K	L	V	F		222
gp190n	A	TC	T	A	T	A	A			C	T	A	AC	T	G		
gp190S	AA	GA	TC	AG	CC	AA	CG	AGT	TG	G	AC	GT	AT	TGA	AGA	AGT	675

THIS PAGE BLANK (USPTO)

AS	G	Y	R	K	P	L	D	N	I	K	D	N	V	G	K	237
gp190n	A	A	A	A	AT	A	T	T	A	T	A	T	A	A	A	
gp190S	GG	AT	AT	CG	CA	GC	CT	CT	CG	CA	CA	AT	CA	AG	GAA	720
AS	M	E	D	Y	I	K	K	N	K	K	T	I	E	N	I	252
gp190n				C				A	A	A	A	A	T	A		
gp190S	AT	GA	AG	AT	TAT	TAA	AA	GA	ATA	AG	AA	GC	AT	CG	AG	765
AS	N	E	L	I	E	E	S	K	K	T	I	D	K	N	K	267
gp190n	T	AT	A	T		AG	T	G	A	A	T	T				
gp190S	A	AC	GA	CT	GA	TC	GA	GA	AT	CC	AA	AA	AG	AC	CA	810
AS	N	A	T	K	E	E	E	K	K	K	L	Y	Q	A	Q	282
gp190n			T	A	A			A	A	A	A	A	A	T	A	
gp190S	A	AT	GA	CA	CA	AG	GA	GA	AA	AA	AA	GA	AG	TT	GT	855
AS	Y	D	L	S	I	Y	N	K	Q	L	E	E	A	H	N	297
gp190n	T	T	T	T	T	C	T	AT	A	A				T		
gp190S	T	AC	CA	CT	CT	AT	AT	CA	CA	CA	AG	CT	TT	GA	AG	900

THIS PAGE BLANK (USPTO)

10/36

AS	L	I	S	V	L	E	K	R	I	D	T	L	K	K	N	312
gp190n	T	A	A	TT	A	A	A	T	T		TT	A	A	A		
gp190S	CTCATCAGCGTACTGGAGAAGCGCATAGACACCCCTCAAGAAGAAT															945
AS	E	N	I	K	E	L	L	D	K	I	N	E	I	K	N	327
gp190n			C	T	G	T	A	T	T	A						
gp190S	GAAATATCAAGAAGAACTGCTCGACAAGATTAAATGAAATTAAGAAT															990
AS	P	P	P	A	N	S	G	N	T	P	N	T	L	L	D	342
gp190n		C	A	G		T		A	T	A	A	T	T	C	T	
gp190S	CCTCCGCCAGCCAACTCTGGGAACACCCCTAACACGCTGCTGGAC															1035
AS	K	N	K	K	I	E	E	H	E	K	E	I	K	E	I	357
gp190n			A	A	C		A		A	A	A	A	A	A	T	
gp190S	AAGAACAAGAAGATAGAGGAGCAGAGAAAGAGATCAAGAGATC															1080
AS	A	K	T	I	K	F	N	I	D	S	L	F	T	D	P	372
gp190n			T		A	T	T	T	AG	T	A				A	
gp190S	GCCAAAACCATTAAGTTCAACATAGATTCTCTCTTTACTGATCCC															1125

THIS PAGE BLANK (USPTO)

AS	L E L E Y Y L R E K N I D	387
gp190n	AT A A T A A A T T	
gp190S	CTTGAGCTGGAGTACTTGTGAGAGAGAATAAGAATATAGAC	1170
AS	I S A K V E T K E S T E P N E	402
gp190n	AAGT A G T A T C	
gp190S	ATCTCCGCCAAAGTCGAGACAAAGGAATCAACCCGAACCTAATGAA	1215
AS	Y P N G V T Y P L S Y N D I N	417
gp190n	A A T T T A T	
gp190S	TATCCCAATGGTGTGACGTACCCCTCTGTCTTATAACGATATCAAC	1260
AS	N A L N E L N S F G D L I N P	432
gp190n	T A T A T TCT T A T A	
gp190S	AACGCTCTCAACGAGCTCAATAGCTTCGGGTGACTTGATTAACCCC	1305
AS	F D Y T K E P S K N I Y T D N	447
gp190n	T A AAG A C A T T T	
gp190S	TTCGATTATACGAAAGAACCCTCTAAGAATATCTACACAGACAAT	1350

THIS PAGE BLANK (USPTO)

AS	E R K K F I N E I K E K I K I	462
gp190n	A A A C A T T A A T A	
gp190S	GAGAGAAAGAAGTTTATCAACGAAATCAAGGAGAAGATCAAAATT	1395
AS	E K K K I E S D K K S Y E D R	477
gp190n	A A A ATC T A TC A A	
gp190S	GAGAAGAAGAAATTGAGAGTGACAAAGAAAGTTACGAAGACCGC	1440
AS	S K S L N D I T K E Y E K L L	492
gp190n	TCT GTC T T A A A AT A T	
gp190S	AGCAAAAGTCTAAACGATATCACTAAAGAGTATGAAAAGCTGCTG	1485
AS	N E I Y D S K F N N N I D L T	507
gp190n	T A T AG T T A TT A T	
gp190S	AACGAGATCTATGATTCCAAATTCAACAATAACATCGACCTGACC	1530
AS	N F E K M M G K R Y S Y K V E	522
gp190n	T A T A A T A T T	
gp190S	AACTTCGAGAAATGATGGGAAAACGGTACTCTTACAAAGTGGAG	1575

THIS PAGE BLANK (USPTO)

AS gp190n K L T H H N T F A S Y E N S K 537
gp190s AACTGACACACCATATAACCTTTGTCATCCTATGAGAAATCTTAAG A 1620

AS gp190n H N L E K L T K A L K Y M E D 552
gp190s CATAATCTTGAGAAGCTCACCAAAGCTCTTAAGTATATGGAGGAC A T 1665

AS gp190n Y S L R N I V V E K E L K Y Y 567
gp190s TATTCTCTGCGGAACATTTGTTGTGGAGAAAGAACTAAAGTATTAC T 1710

AS gp190n K N L I S K I E N E I E T L V 582
gp190s AAGAATCTCATTAAGTAAGATCGAAACGAGATCGAGACGCTTGTT T 1755

AS gp190n E N I K K D E E Q L F E K K I 597
gp190s GAGAACATTAAGAAGGATGAAGAACAGTTGTTTGAGAAAGAGATT A A A A 1800

THIS PAGE BLANK (USPTO)

AS	T	K	D	E	N	K	P	D	E	K	I	L	E	V	S	612
gp190n	T									A	A	TT	A	A	A	T
gp190S	ACAAAAGACGAAATAAACCCAGATGAGAAAGATCCTGGAGGCTCTCC															1845
AS	D	I	V	K	V	Q	V	Q	K	V	L	L	M	N	K	627
gp190n	C	A	A	T	A	A	TT	AT	A						A	
gp190S	GATATTGTTAAAGTCCAAGTGCAGAAAGGTGCTCCTCATGAACAAG															1890
AS	I	D	E	L	K	K	T	Q	L	I	L	K	N	V	E	642
gp190n	C	T	A	A			T	G	T	A	A	T	A	A		
gp190S	ATTGATGAACTCAAGAAGACTCAACTCATTTCTGAAGAACGTGGAG															1935
AS	L	K	H	N	I	H	V	P	N	S	Y	K	Q	E	N	657
gp190n							T	C	TC	C	A	A	A			
gp190S	TTAAACATAATATACATGTGCCGAATAGTTATAAGCAGGAGAAT															1980
AS	K	Q	E	P	Y	Y	L	I	V	L	K	K	E	I	D	672
gp190n	A	T	T	T	A	T	GT	G	A	A	T	T				
gp190S	AAGCAGGAACCATACTACCTCATCTCGTACTCAAGAAAGAGATAGAC															2025

THIS PAGE BLANK (USPTO)

15 / 36

AS	K	L	K	V	F	M	P	K	V	E	S	L	I	N	E	687
gp190n	T	A						T	G	A	ATCAT			A	T	
gp190S	AAACTGAAAGTGTTCATGCCCAAGTCGAGAGCCTGATCAACGAA															2070
AS	E	K	K	N	I	K	T	E	G	Q	S	D	N	S	E	702
gp190n	A	A	A		A		A		T	A	G		T	G	A	
gp190S	GAGAAGAAGAACATTAAAACTGAAGGACAGTCAGATAACTCCGAG															2115
AS	P	S	T	E	G	E	I	T	G	Q	A	T	T	K	P	717
gp190n	A	A	C			A		A		A	A	T	A	A	T	
gp190S	CCTTCCACAGAAGGAGAGAGATAACCGGACAGGCTACCAACCAAGCCC															2160
AS	G	Q	Q	A	G	S	A	L	E	G	D	S	V	Q	A	732
gp190n	A	A	A	A	T	T	A		A		TCA		A		A	
gp190S	GGACAACAGGCCGGTTCAGCTCTCGAAGGCCGATAGCGTGCAAGCT															2205
AS	Q	A	Q	E	Q	K	Q	A	Q	P	P	V	P	V	P	747
gp190n				A	A	A	A		A	A		A		A	A	
gp190S	CAAGCACAAAGAGCAGAAGCAGGCACAGCCTCCAGTGCCAGTGCCC															2250

THIS PAGE BLANK (USPTO)

16 / 36

AS	V	P	E	A	K	A	Q	V	P	T	P	P	A	P	V	762	
gp190n	A	A	A	A	A	A	C	A					A	A	A		
gp190S	G	T	T	C	A	G	A	G	T	C	C	T	A	C	A	C	2295
AS	N	N	K	T	E	N	V	S	K	L	D	Y	L	E	K	777	
gp190n	T	A	T	A		T	T	C		T	A	T	T	A	A		
gp190S	A	A	T	A	C	C	G	A	A	T	G	T	C	A	A	C	2340
AS	L	Y	E	F	L	N	T	S	Y	I	C	H	K	Y	I	792	
gp190n	T	A	A	T	T	A	T	A	T	A	T				T		
gp190S	C	T	C	T	A	T	G	A	T	A	C	T	C	T	A	C	2385
AS	L	V	S	H	S	T	M	N	E	K	I	L	K	Q	Y	807	
gp190n	T	G	T	A	T	C	A		A		A	T	A	A	T		
gp190S	C	T	C	T	C	T	A	C	A	C	T	A	T	G	A	A	2430
AS	K	I	T	K	E	E	S	K	L	S	S	C	D	P		822	
gp190n	A	T	A		G	A	A	C	T	A	A	G	T	A			
gp190S	A	A	G	A	T	A	C	G	A	G	A	G	T	A	A	A	2475

THIS PAGE BLANK (USPTO)

17/36

AS	L D L L F N I Q N N I P V M Y	837
gp190n	T A T A T T A A T A T A	
gp190S	CTGGACCTGCTGTTC AATATCCAGAACACAATTC CCGTTATGTAT	2520
AS	S M F D S L N N S L S Q L F M	852
gp190n	T T A T A A G T A A A T	
gp190S	TCTATGTTTCGATAGCCCTCAACAATTC TCTCTCTCAACTGTTCATG	2565
AS	E I Y E K E M V C N L Y K L K	867
gp190n	A T A A A T T T T A T G	
gp190S	GAGATATATGAGAAGGAGATGGTCTGCAACCTGTATAAACTCAA	2610
AS	D N D K I K N L L E E A K K V	882
gp190n	T T A A A T T A T A G A A A	
gp190S	GACAACGACAAGATTAAAGAACCTTCTGAGGAAGCTAAGAAGGTC	2655
AS	S T S V K T L S S S S M Q P L	897
gp190n	A A T A A G T T C A A T A	
gp190S	TCCACCTCTGTTA A A A C T C T C T C T C C A G C T C C A T G C A A C C A C T G	2700

THIS PAGE BLANK (USPTO)

AS S L T P Q D K P E V S A N D D 912
gp190n AT A G T A A T A T T T
gp190s TCTCTCACACCTCAAGACAAGCCGAAGTGAGCGCTAACGACGAC 2745

AS T S H S T N L N N S L K L F E 927
gp190n A A T T A T T G TAGTT A T A A
gp190s ACCTCTCACTCGACCAACCTTAATAACTCACTGAAACTGTTTGAG 2790

AS N I L S L G K N K N I Y Q E L 942
gp190n AT AG T A A C A T A T A
gp190s AACATCCTGTCTCTCGCAAGAATAAGAACATCTACCAAGAACTT 2835

AS I G Q K S S E N F Y E K I L K 957
gp190n A T A AGTAGT A T T A T A
gp190s ATTGGACAGAAATCGTCCGAGAACTTCTACGAGAAGATACTGAAA 2880

AS D S D T F Y N E S F T N F V K 972
gp190n T T T T ATCT T A T T A
gp190s GACAGCGACACATTCTATATAACGAGAGCTTCACTAACTTCGTGAAA 2925

THIS PAGE BLANK (USPTO)

AS S K A D D I N S L N D E S K R 987
gp190n T T T A T G T A A G
gp190s TCTAAAGCCGATGATATCAACTCTCTTAACGATGAATCTAAACGT 2970

AS K K L E E D I N K L K K T L Q 1002
gp190n AT A A T T AT A A A TT A G
gp190s AAGAAGCTGGAAGAGGACATCAATAAGCTGAAGAAGACACTGCAA 3015

AS L S F D L Y N K Y K L K L E R 1017
gp190n T ATCA T TT A T T A T T A T A A
gp190s CTGAGCTTCGACCCTGTACAACAAGTACAAACTGAAACTGGAGAGA 3060

AS L F D K K K T V G K Y K M Q I 1032
gp190n T A T T A A T T A A A A T
gp190s CTCTTCGACAAGAAGAAGACAGTCGGCAAGTATAAGATGCAGATC 3105

AS K K L L L L K E Q L E S K L N 1047
gp190n A AC T T AT A A A AT A TCA T G T
gp190s AAGAAGTTGACTCTGCTCAAGGAGCAGCTTGAAAGCAAACCTCAAC 3150

THIS PAGE BLANK (USPTO)

AS S L N N P K H V L Q N F S V F 1062
gp190n T T C A G T T A A T T T T
gp190s TCACTGAACAATCCGAAACACGTA CTGCAGAACTTCTCAGTG TTC 3195

AS F N K K K E A E I A E T E N T 1077
gp190n T A A A T A A A A T A A
gp190s TTCAACAAGAAGGAGCCGAGATCGCCGAGACAGAGA CACT 3240

AS L E N T K I L L K H Y K G L V 1092
gp190n T A A A A T A T G T T A T T
gp190s CTGGAGAACACCAAGATTCTTCTCAAAACACTACAAAGGCC TCCTCGTC 3285

AS K Y Y N G E S S P L K T L S E 1107
gp190n A T A A AT A A T AAGT A
gp190s AAGTATTATAATGGCGAGTCTTCTCTGAAGACTCTCTCCGAG 3330

AS E S I Q T E D N Y A S L E N F 1122
gp190n ATCA T A A A T T T T A A T
gp190s GAGAGCATCCAGACCGAGGATAACTACGCCAGCCTCGAGAACTTC 3375

THIS PAGE BLANK (USPTO)

21/36

AS	K	V	L	S	K	L	E	G	K	L	K	D	N	L	N	1137
gp190n	A	AT	AAG	AT	A	A	AT	A	T	TT	A	T				
gp190s	AAG	TCCT	GTCT	TAAG	CTCG	AAGG	CAAG	CTGA	AGGACA	ACCT	GAAC					3420
AS	L	E	K	K	L	S	Y	L	S	S	G	L	H	H		1152
gp190n	T	A	A	A	AT	ATCA	T	A	A	T	T	A	T			
gp190s	CTG	GAGA	AGAAG	CTCAG	CTAC	CTCTCT	CTAG	CGGACT	GCAT	CAC						3465
AS	L	I	A	E	L	K	E	V	I	K	N	K	N	Y	T	1167
gp190n	T	A	T	T	AT	A	A	A	A	T	A	T	T	A		
gp190s	CTG	ATCG	CCGAG	CTCA	AGGA	AGTCAT	TAA	GAA	CAAG	AACT	ACAC					3510
AS	G	N	S	P	S	E	N	N	T	D	V	N	N	A	L	1182
gp190n	T	TCT	T	A	G	T	T	C	T	T	A					
gp190s	GGC	AATAG	CCCAAG	CGAGA	ATAATA	CAGACG	TGA	ATAAC	GCACTG							3555
AS	E	S	Y	K	K	F	L	P	E	G	T	D	V	A	T	1197
gp190n	A	A	T	C	A	T	A	A								
gp190s	GA	ATCT	TACA	GAAG	TTCT	CGCTG	CAAG	GAAC	AGAT	GT	CGCC	ACT				3600

THIS PAGE BLANK (USPTO)

22 / 36

AS	V	V	S	E	S	G	S	D	T	L	E	Q	S	Q	P	1212
gp190n	T	AAG	AG	A					T	A	A	AAG			A	
gp190S	GTGGTGTCTGAATCTGGCTCCGACACACTGGAGCAGTCTCAACCT															3645
AS	K	K	P	A	S	T	H	V	G	A	E	S	N	T	I	1227
gp190n	A	A	A	A				A	A			T	C		A	
gp190S	AAGAAGCCTGCATCTACTCTCATGTCTGGAGCCGAGTCCAATAAATT															3690
AS	T	T	S	Q	N	V	D	D	E	V	D	D	V	I	I	1242
gp190n	A	A	A	T	T			A	A			A			A	
gp190S	ACCACATCTCAGAACGTCGACGATGAGGTCGATGACGTCATCATTT															3735
AS	V	P	I	F	G	E	S	E	E	D	Y	D	D	L	G	1257
gp190n	A	A	T	A	ATC	A	A	T	T			TT	A		A	
gp190S	GTGCCCTATCTTCGGCGAGAGCGAGGAGGACTACGATGACCTCGGC															3780
AS	Q	V	V	T	G	E	A	V	T	P	S	V	I	D	N	1272
gp190n	A	A	A	A	A	A	A								A	
gp190S	CAGGTGTCACCGGTGAGGCTGTCACTCCTTCCGTGATTGATAAC															3825

THIS PAGE BLANK (USPTO)

```

AS      I L S K I E N E Y E V L Y L K      1287
gp190n  A T T T A T T G T T A T A
gp190S  ATTCTGTCCAAATCGAGAACGAATACGAAGTGCTCTATCTGAAA 3870

AS      P L A G V Y R S L K K Q L E N      1302
gp190n  T A T T AAG T A A AT A A
gp190S  CCTCTGGCAGCGTCTATAGGTCTCTCAAGAAACAGCTGGAGAAT 3915

AS      N V M T F N V N V K D I L N S      1317
gp190n  T A T T T T T T T A TTCA
gp190S  AACGTGATGACCTTCAATGTCAACGTGAAGGACATTTCTGAACAGC 3960

AS      R F N K R E N F K N V L E S D      1332
gp190n  A AC T A T T A ATCA T
gp190S  CGCTTTAATAAGAGAGAGAAAATTTCAAGAACGTCTTTGGAGAGCGAC 4005

AS      L I P Y K D L T S S N Y V V K      1347
gp190n  A A TT A A AAG T T A
gp190S  TTGATTCCCTATAAAGACCTGACCTCCTCTAACTACGTTGTCAAG 4050

```

THIS PAGE BLANK (USPTO)

24 / 36

AS	D	P	Y	K	F	L	N	K	E	K	R	D	K	F	L	1362
gp190n	T	T	A	T	T		A	A	A						CT A	
gp190s	G	A	C	C	A	T	A	A	G	T	T	C	A	A	G	4095
AS	S	S	Y	N	Y	I	K	D	S	I	D	T	D	I	N	1377
gp190n	A	G	C	T	T	T	T	A	A	T	G				A	
gp190s	T	C	T	A	G	T	T	A	C	A	G	A	C	T	C	4140
AS	F	A	N	D	V	L	G	Y	Y	K	I	L	S	E	K	1392
gp190n	T	A		T	T	A		T	A	T	A	A	T	A	T	
gp190s	T	T	C	G	T	A	A	T	T	A	C	A	A	G	A	4185
AS	Y	K	S	D	L	D	S	I	K	K	Y	I	N	D	K	1407
gp190n	T	A	A	T	T	A	T	A		A					C A	
gp190s	T	A	C	A	A	G	T	C	T	A	T	A	A	A	A	4230
AS	Q	G	E	N	E	K	Y	L	P	F	L	N	N	I	E	1422
gp190n	T	A		G		C	T			T	T	A	C	T	T G	
gp190s	C	A	A	G	G	A	A	T	A	T	C	T	C	T	G	4275

THIS PAGE BLANK (USPTO)

AS	T	L	Y	K	T	V	N	D	K	I	D	L	F	V	I	1437
gp190n	T	A	T	A	T	T	T	T	T	T	T	T	T	A	T	
gp190s	AC	CT	GT	AC	AG	AG	CA	CA	AA	AT	CG	AC	CT	CT	TC	4320
AS	H	L	E	A	K	V	L	N	Y	T	Y	E	K	S	N	1452
gp190n	TT	A	A	A	A	T	A	T	A	T	A	T	AT	CA	C	
gp190s	CAC	CT	GG	AG	CC	AG	TC	CT	CA	ACT	AT	ACT	TAC	GAG	AA	4365
AS	V	E	V	K	I	K	E	L	N	Y	L	K	T	I	Q	1467
gp190n	A			A	A	A	T	T	T	T	A			T		
gp190s	GT	GA	AG	TT	AA	AT	CA	AG	AG	CT	GA	ACT	TAC	CT	CA	4410
AS	D	K	L	A	D	F	K	K	N	N	N	F	V	G	I	1482
gp190n	AT						T	A						T		
gp190s	GAC	A	AG	CT	GC	AG	AT	TT	CA	AG	AA	AA	TA	ACA	AT	4455
AS	A	D	L	S	T	D	Y	N	H	N	N	L	L	T	K	1497
gp190n	T	TT	A	A	A				T	T	CT	AT		A		
gp190s	GC	AG	AC	CT	GT	CT	AC	CG	AT	TAT	AAC	CA	CA	CA	AT	4500

THIS PAGE BLANK (USPTO)

AS	F	L	S	T	G	M	V	F	E	N	L	A	K	T	V	1512
gp190n	C	T	A	G	T	A	T	T	T	T	T	T	T	C	T	
gp190S	T	T	T	C	T	C	A	T	G	G	T	G	T	T	C	4545
AS	L	S	N	L	L	D	G	N	L	Q	G	M	L	N	I	1527
gp190n	T	A	T	C	T	T	A	T	A	T	A	T	T	A	T	
gp190S	C	T	G	A	G	C	A	A	C	T	G	C	A	G	G	4590
AS	S	Q	H	Q	C	V	K	K	Q	C	P	Q	N	S	G	1542
gp190n	A	A				A	A	A	A	T	A	A	A	T	C	
gp190S	T	C	C	A	G	C	A	A	T	G	C	G	T	G	A	4635
AS	C	F	R	H	L	D	E	R	E	E	C	K	C	L	L	1557
gp190n																
gp190S	T	G	T	T	C	A	G	G	C	A	T	C	T	G	C	4680
AS	N	Y	K	Q	E	G	D	K	C	V	E	N	P	N	P	1572
gp190n	T															
gp190S	A	A	T	A	A	G	A	T	A	A	G	T	G	G	A	4725

THIS PAGE BLANK (USPTO)

```

AS      T C N E N N G G C D A D A K C      1587
gp190n  T T C T T A T A C T      T
gp190s  ACCTGCAATGAAACAATGGCGGTGTGACCGCGATGCTAAATGC 4770

AS      T E E D S G S N G K K I T C E      1602
gp190n  A TTCA TAGC T A
gp190s  ACCGAGGAAGACAGCGGCTCTAACGGAAGAAATCACATGCGAG 4815

AS      C T K P D S Y P L F D G I F C      1617
gp190n  A T T T T T T T C
gp190s  TGTACTAAGCCCGACTCCTATCCACTCTTCGACGGGATTTTTCG 4860

AS      S S S N F L G I F F L L I L M      1632
gp190n  AGTTC C T A A A CA T AT A A
gp190s  TCCAGCTCTAATTTCCTGGGCATCTTCTTCCCTGCTGATCCTCATG 4905

AS      L I L Y S F I * *      1639
gp190n  T A AT A T T
gp190s  CTGATCCTGTACAGCTTCATCTAATAGATCGATGG 4940
              stop codon Cla I

```

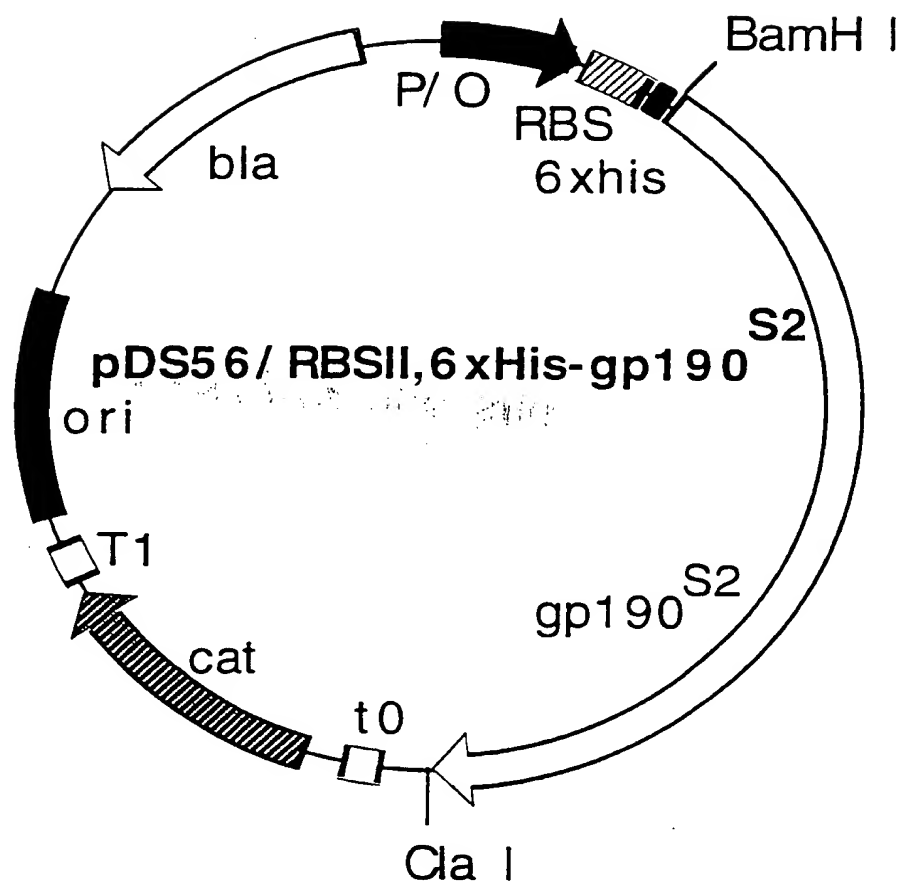
THIS PAGE BLANK (USPTO)

Abb.3D

	N'-terminus		C'-terminus	
gp190s1 Sequence				
DNA Sequence	GC <u>ACGCGTATGAAAATC</u> -----		AGCTCTAAATTAATAGGCGGGCCGCATCGATGGC	
AA Sequence	Mlu I Met Lys Ile -----		Ser Ser Asn stop codon Not I Cla I	
AA Position	1 2 3 -----		1619 1620 1621	
gp190s2 Sequence				
DNA Sequence	GC <u>GGATCCGTGACCCAC</u> -----		AGCTCTAAATTAATAGGCGGGCCGCATCGATGGC	
AA Sequence	BamH I Val Thr His -----		Ser Ser Asn stop codon Not I Cla I	
AA Position	20 21 22 -----		1619 1620 1621	

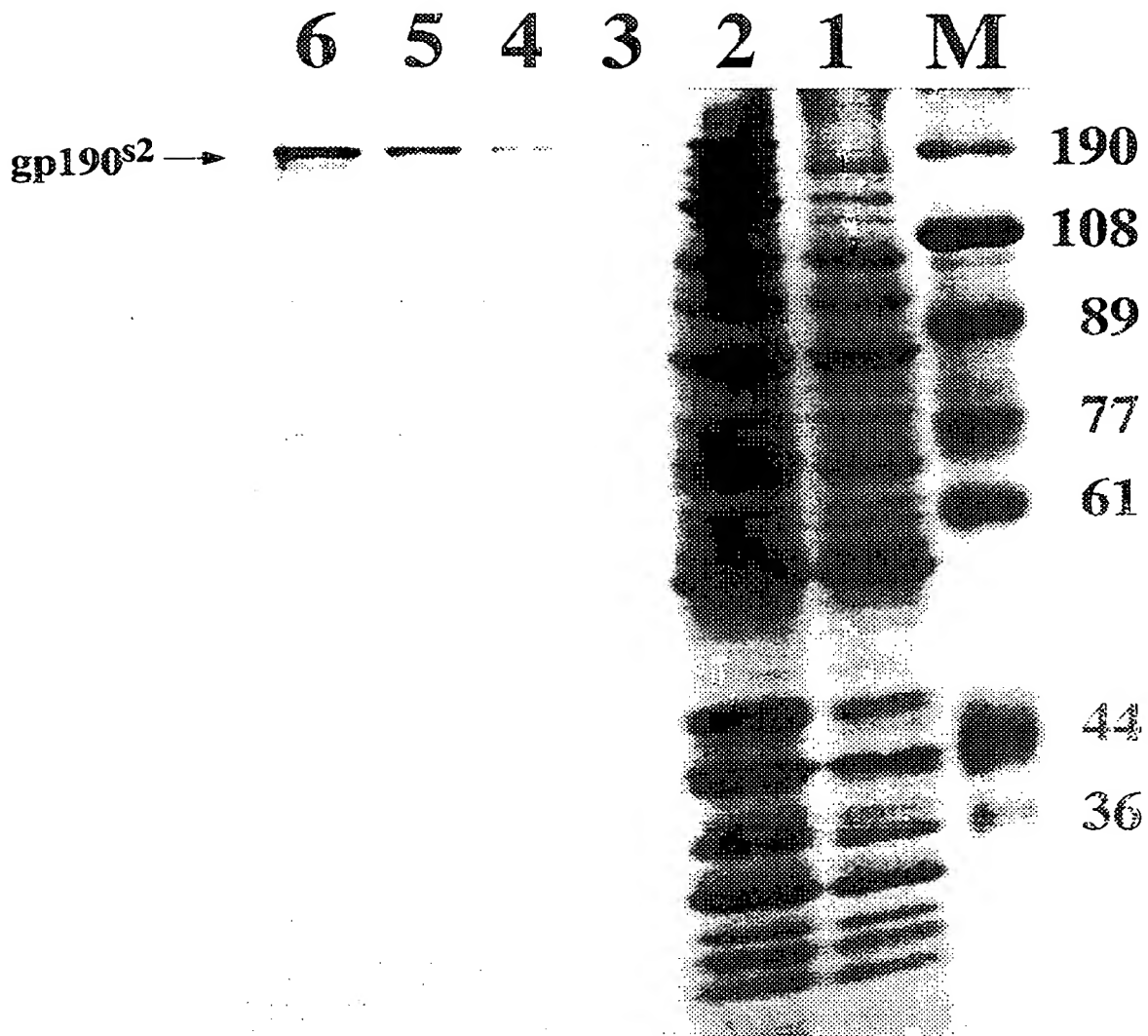
THIS PAGE BLANK (USPTO)

Abb.4 A



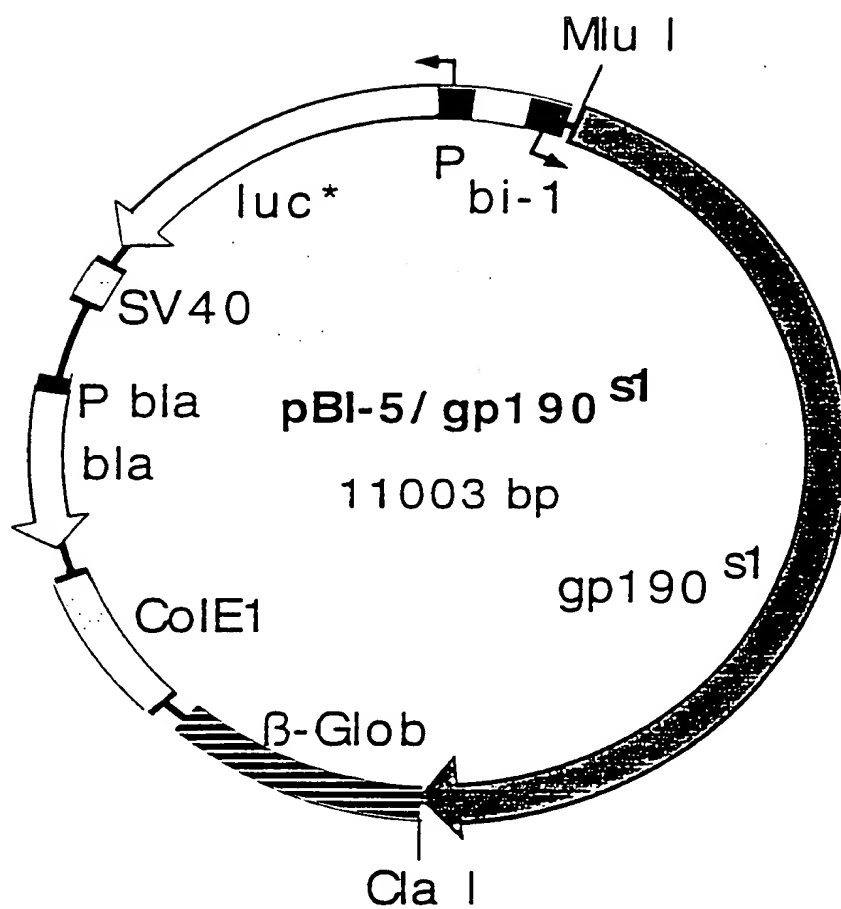
THIS PAGE BLANK (USPTO)

Fig. 4B



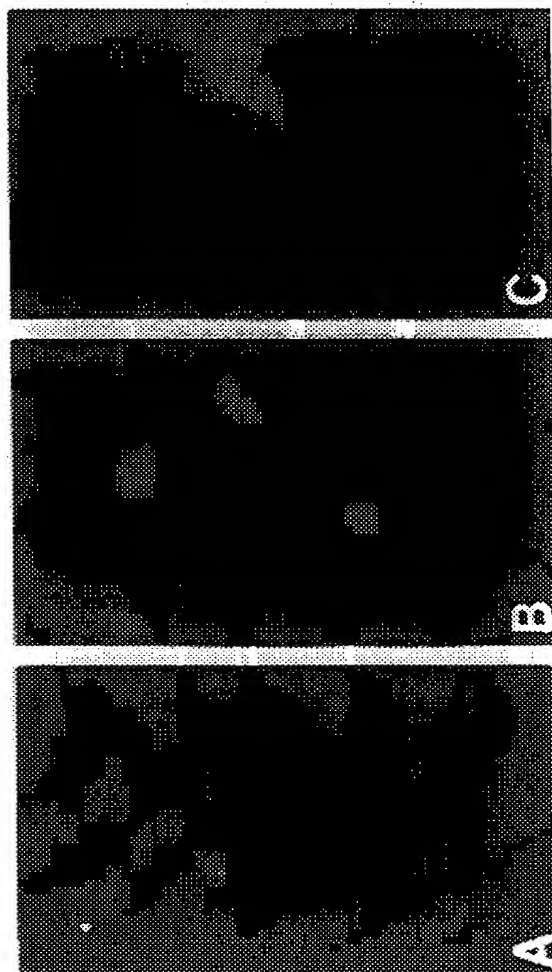
THIS PAGE BLANK (USPTO)

Abb.5 A

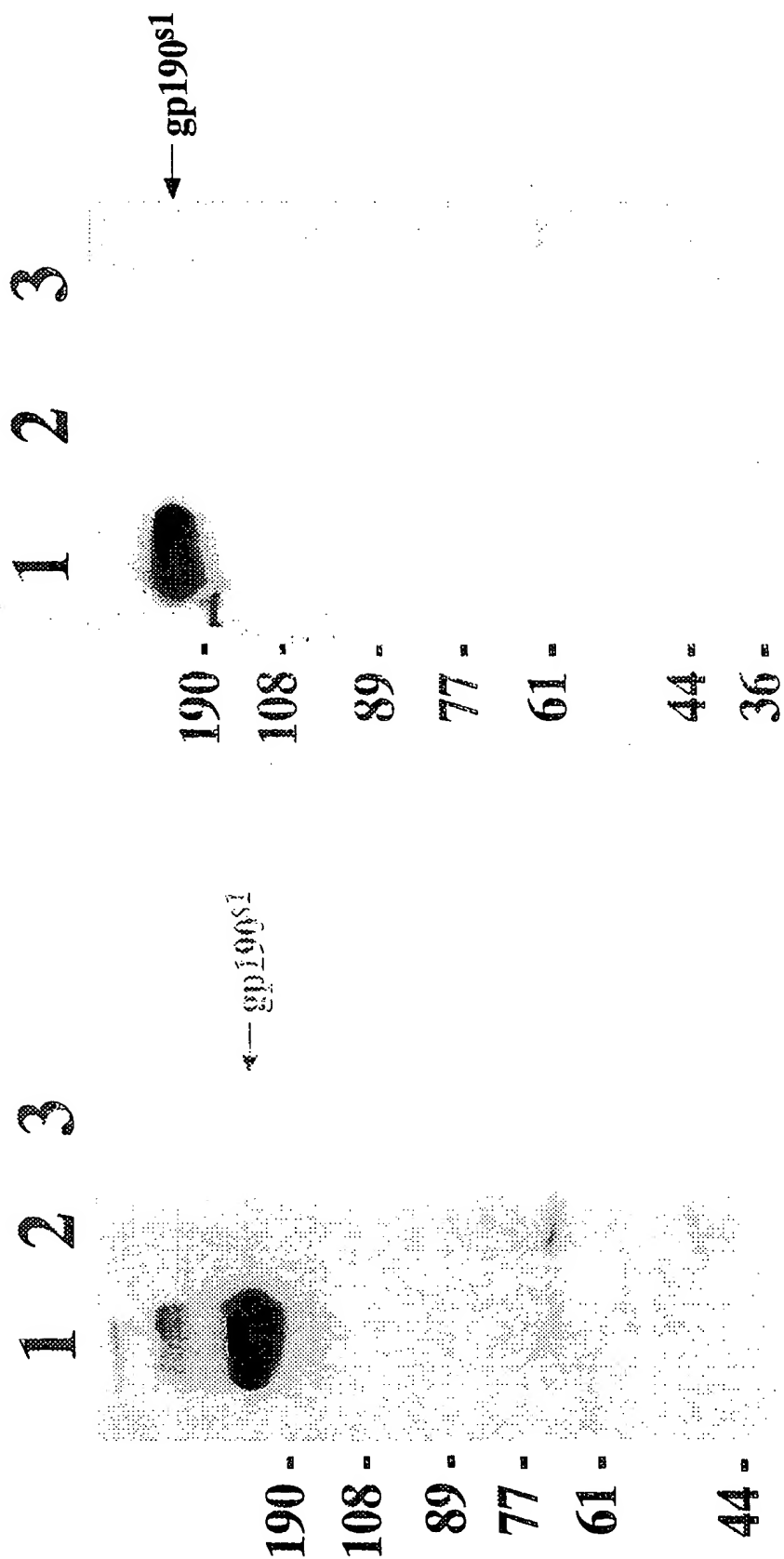


THIS PAGE BLANK (USPIC)

Fig. 5B



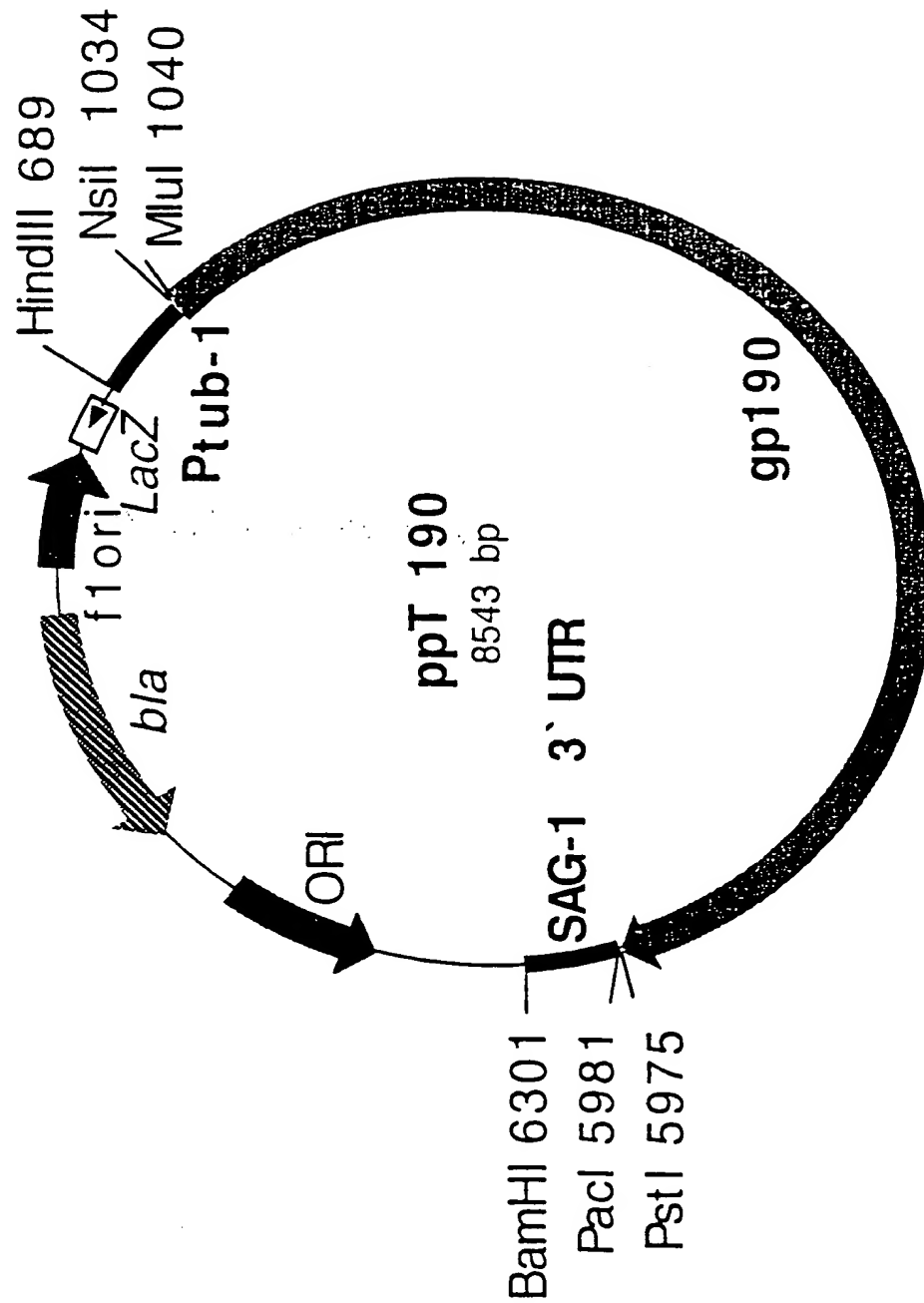
THIS PAGE BLANK (USPTO)



THIS PAGE BLANK (USPTO)

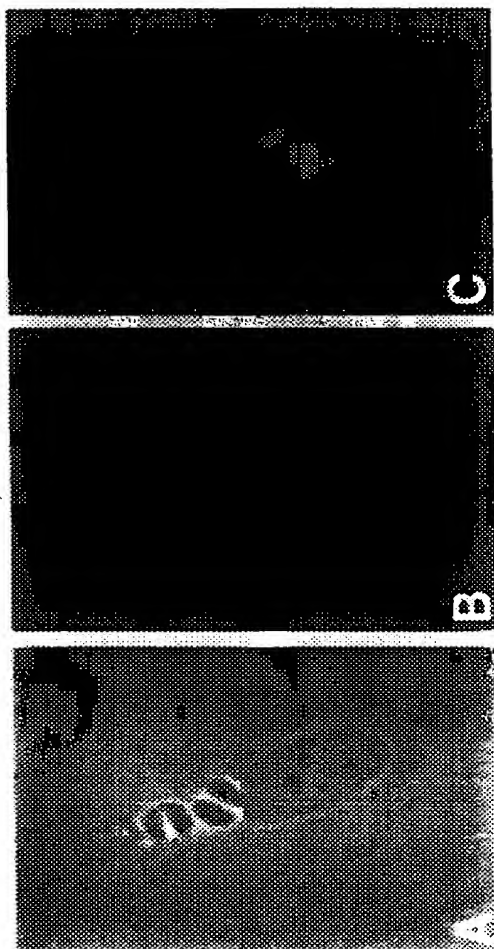
34 / 36

Abb. 6A



THIS PAGE BLANK (USPTO)

Fig. 6B



THIS PAGE BLANK (USPTO)

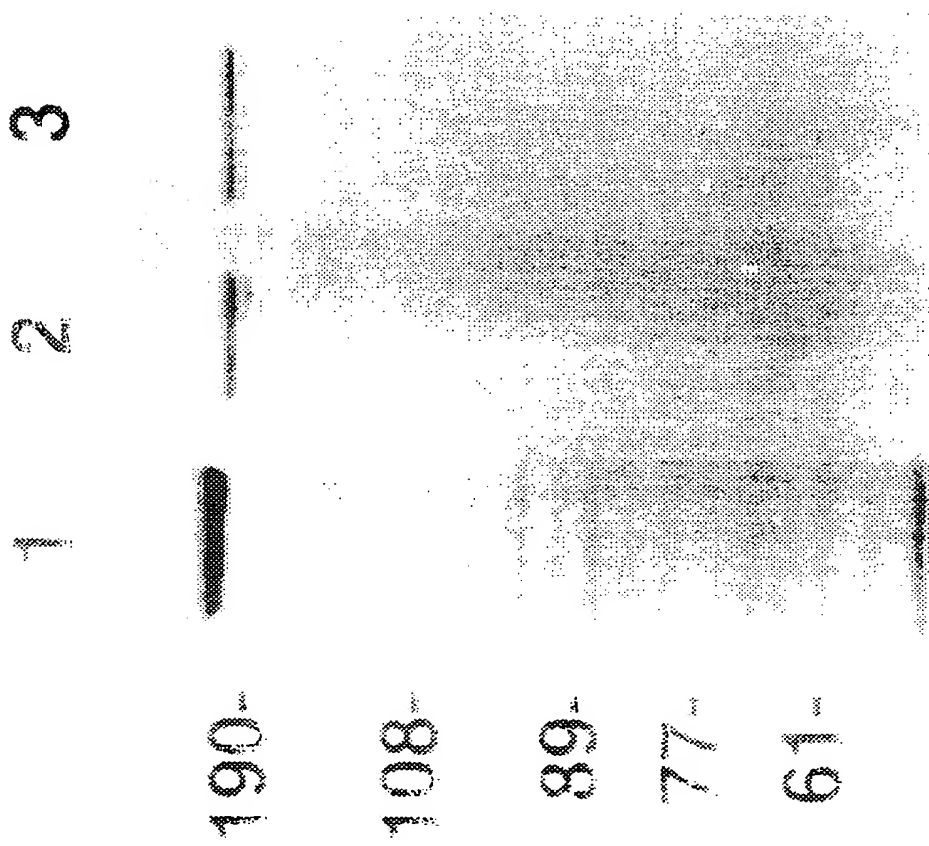


Fig. 6c

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

Intern. Patent Application No.
PCT/EP 97/05441

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/30 C07K14/445 C12N15/62 A61K39/015 A61K31/70
C07H21/00 C12N1/21 C12N5/10 C12N1/11 C12N15/67

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 28930 A (VIROGENETICS CORP) 22 December 1994	1-4, 12, 14, 17, 26-29, 33, 34, 39, 40, 18, 20
Y	see page 19 - page 22; claims 1-33; examples 5, 29, 53, 57, 63	
Y	EP 0 340 359 A (WELLCOME FOUND) 8 November 1989 see figure 2; examples 3, 4	18

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

6 May 1998

Date of mailing of the international search report

18.05.98

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Espen, J

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 97/05441

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KASLOW DC ET AL: "Expression and antigenicity of Plasmodium falciparum major merozoite surface protein (MSP1(19)) variants secreted from Saccharomyces cerevisiae." MOL BIOCHEM PARASITOL, FEB 1994, 63 (2) P283-9, NETHERLANDS, XP000603953 see the whole document	20
X	--- HOLDER A A ET AL: "PRIMARY STRUCTURE OF THE PRECURSOR TO THE THREE MAJOR SURFACE ANTIGENS OF PLASMODIUM FALCIPARUM MEROZOITES" NATURE, vol. 317, 19 September 1985, pages 270-273, XP000604859 see figure 2	17
X	--- MYLER P J: "Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plasmodium falciparum Palo Alto PLF-3/B11" NUCLEIC ACIDS RESEARCH., vol. 17, no. 13, 1989, OXFORD GB, page 5401 XP002057620 see the whole document	17
X	--- EP 0 154 454 A (WELLCOME FOUND) 11 September 1985 see the whole document	1,12,13, 17, 26-28, 39,40
X	--- SIDDIQUI W A ET AL: "Merozoite surface coat precursor protein completely protects Aotus monkeys against Plasmodium falciparum malaria" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., vol. 84, 1987, WASHINGTON US, pages 3014-3018, XP002057621 cited in the application see the whole document	33-35, 39,40
X	--- GENTZ R ET AL: "Major surface antigen p190 of Plasmodium falciparum: detection of common epitopes present in a variety of plasmodia isolates" EMBO JOURNAL., vol. 7, no. 1, 1988, EYNSHAM, OXFORD GB, pages 225-230, XP002057622 see figure 1; table 1 --- -/--	17-28, 33,34, 39,40

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 97/05441

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PAN W ET AL: "A direct and rapid sequencing strategy for the Plasmodium falciparum antigen gene gp190/MSA1." MOL BIOCHEM PARASITOL, JUL 1995, 73 (1-2) P241-4, NETHERLANDS, XP002057623 see page 1; table 1 ---	17
X	EP 0 385 962 A (MONSANTO CO) 5 September 1990 see figures 2-4; examples 2,3 ---	36-40
X	EP 0 359 472 A (LUBRIZOL GENETICS INC) 21 March 1990 see page 9; figure 1; table 3 see page 22, line 1 - line 33 -----	36-40

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP 97/05441

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Observation: Although Claims 33-35 relate to a method for treatment of the human/animal body, the search was carried out and was based on the cited effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See Supplemental Sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

R mark n Pr test

☐
☒

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

Claims: 1-35, partially 37-40

DNA sequence coding gp190/MSP-1 plasmodium surface protein, host organism containing said sequence, use of gp190/MSP-1 plasmodium surface protein or the therefor coding DNA for immunization against malaria, vector containing said DNA sequence, vaccine containing gp190/MSP-1 plasmodium surface protein or the therefor coding DNA, method for the production of gp190/MSP-1 plasmodium surface protein

2.Claims: 36, partially, 37-40

Method for stabilizing gene sequences by reducing the AT content of the sequences, stabilized gene, vector containing said gene, vaccine containing said vector.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter. onal Application No

PCT/EP 97/05441

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9428930 A	22-12-94	AU 7060294 A EP 0717636 A	03-01-95 26-06-96
EP 0340359 A	08-11-89	AU 620041 B AU 2156988 A CA 1331155 A DE 3882522 A DE 3882522 T DK 478588 A ES 2058294 T IE 62490 B JP 2115464 C JP 2167088 A JP 8013275 B PT 88362 A,B US 5147788 A	13-02-92 09-11-89 02-08-94 26-08-93 03-03-94 07-11-89 01-11-94 08-02-95 06-12-96 27-06-90 14-02-96 30-11-89 15-09-92
EP 0154454 A	11-09-85	AU 592749 B AU 3904685 A DK 79985 A GB 2154592 A IL 74409 A JP 2584733 B JP 61019490 A JP 6189772 A PH 25993 A US 5597708 A	25-01-90 05-09-85 23-08-85 11-09-85 26-08-94 26-02-97 28-01-86 12-07-94 13-01-92 28-01-97
EP 0385962 A	05-09-90	AU 638438 B AU 5163090 A CA 2024811 A EP 0413019 A IL 93513 A JP 3504333 T WO 9010076 A US 5500365 A	01-07-93 26-09-90 25-08-90 20-02-91 14-11-95 26-09-91 07-09-90 19-03-96
EP 0359472 A	21-03-90	AT 132193 T AU 623429 B AU 4118289 A	15-01-96 14-05-92 15-03-90

INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/EP 97/05441

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0359472 A		CN 1044298 A	01-08-90
		DE 68925253 D	08-02-96
		EP 0682115 A	15-11-95
		ES 2083384 T	16-04-96
		JP 2186989 A	23-07-90
		US 5380831 A	10-01-95
		US 5567600 A	22-10-96
		US 5567862 A	22-10-96

THIS PAGE BLANK (USPTO)

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 97/05441

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES

IPK 6 C12N15/30 C07K14/445 C12N15/62 A61K39/015 A61K31/70
C07H21/00 C12N1/21 C12N5/10 C12N1/11 C12N15/67

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 6 C12N C07K A61K C07H

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	WO 94 28930 A (VIROGENETICS CORP) 22.Dezember 1994	1-4,12, 14,17, 26-29, 33,34, 39,40 18,20
Y	siehe Seite 19 - Seite 22; Ansprüche 1-33; Beispiele 5,29,53,57,63	
Y	EP 0 340 359 A (WELLCOME FOUND) 8.November 1989 siehe Abbildung 2; Beispiele 3,4	18
	-/--	

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

☒ Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E Älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

Z Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

6.Mai 1998

Absenddatum des internationalen Recherchenberichts

18.05.98

Name und Postanschrift der Internationalen Recherchenbehörde
Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Espen, J

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	KASLOW DC ET AL: "Expression and antigenicity of Plasmodium falciparum major merozoite surface protein (MSP1(19)) variants secreted from Saccharomyces cerevisiae." MOL BIOCHEM PARASITOL, FEB 1994, 63 (2) P283-9, NETHERLANDS, XP000603953 siehe das ganze Dokument	20
X	--- HOLDER A A ET AL: "PRIMARY STRUCTURE OF THE PRECURSOR TO THE THREE MAJOR SURFACE ANTIGENS OF PLASMODIUM FALCIPARUM MEROZOITES" NATURE, Bd. 317, 19.September 1985, Seiten 270-273, XP000604859 siehe Abbildung 2	17
X	--- MYLER P J: "Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plasmodium falciparum Palo Alto PLF-3/B11" NUCLEIC ACIDS RESEARCH., Bd. 17, Nr. 13, 1989, OXFORD GB, Seite 5401 XP002057620 siehe das ganze Dokument	17
X	--- EP 0 154 454 A (WELLCOME FOUND) 11.September 1985 siehe das ganze Dokument	1,12,13, 17, 26-28, 39,40
X	--- SIDDIQI W A ET AL: "Merozoite surface coat precursor protein completely protects Aotus monkeys against Plasmodium falciparum malaria" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., Bd. 84, 1987, WASHINGTON US, Seiten 3014-3018, XP002057621 in der Anmeldung erwähnt siehe das ganze Dokument	33-35, 39,40
X	--- GENTZ R ET AL: "Major surface antigen p190 of Plasmodium falciparum: detection of common epitopes present in a variety of plasmodia isolates" EMBO JOURNAL., Bd. 7, Nr. 1, 1988, EYNSHAM, OXFORD GB, Seiten 225-230, XP002057622 siehe Abbildung 1; Tabelle 1 --- -/--	17-28, 33,34, 39,40

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	PAN W ET AL: "A direct and rapid sequencing strategy for the Plasmodium falciparum antigen gene gp190/MSA1." MOL BIOCHEM PARASITOL, JUL 1995, 73 (1-2) P241-4, NETHERLANDS, XP002057623 siehe Seite 1; Tabelle 1 ---	17
X	EP 0 385 962 A (MONSANTO CO) 5.September 1990 siehe Abbildungen 2-4; Beispiele 2,3 ---	36-40
X	EP 0 359 472 A (LUBRIZOL GENETICS INC) 21.März 1990 siehe Seite 9; Abbildung 1; Tabelle 3 siehe Seite 22, Zeile 1 - Zeile 33 -----	36-40

INTERNATIONALER RECHERCHENBERICHT

ationales Aktenzeichen

PCT/EP 97/05441

Feld I Bemerkungen zu den Ansprüchen, die sich als nicht recherchierbar erwiesen haben (Fortsetzung von Punkt 1 auf Blatt 1)

Gemäß Artikel 17(2)a) wurde aus folgenden Gründen für bestimmte Ansprüche kein Recherchenbericht erstellt:

1. ☒ Ansprüche Nr.
weil Sie sich auf Gegenstände beziehen, zu deren Recherche die Behörde nicht verpflichtet ist, nämlich
Bemerkung : Obwohl die Ansprüche 33-35 sich auf ein Verfahren zur Behandlung des menschlichen/tierischen Körpers beziehen, wurde die Recherche durchgeführt und gründete sich auf die angeführten Wirkungen der Verbindung/Zusammensetzung.
2. ☐ Ansprüche Nr.
weil sie sich auf Teile der internationalen Anmeldung beziehen, die den vorgeschriebenen Anforderungen so wenig entsprechen, daß eine sinnvolle internationale Recherche nicht durchgeführt werden kann, nämlich
3. ☐ Ansprüche Nr.
weil es sich dabei um abhängige Ansprüche handelt, die nicht entsprechend Satz 2 und 3 der Regel 6.4 a) abgefaßt sind.

Feld II Bemerkungen bei mangelnder Einheitlichkeit der Erfindung (Fortsetzung von Punkt 2 auf Blatt 1)

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere Erfindungen enthält:

siehe Zusatzblatt

1. ☒ Da der Anmelder alle erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht auf alle recherchierbaren Ansprüche der internationalen Anmeldung.
2. ☐ Da für alle recherchierbaren Ansprüche die Recherche ohne einen Arbeitsaufwand durchgeführt werden konnte, der eine zusätzliche Recherchegebühr gerechtfertigt hätte, hat die internationale Recherchenbehörde nicht zur Zahlung einer solchen Gebühr aufgefordert.
3. ☐ Da der Anmelder nur einige der erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht nur auf die Ansprüche der internationalen Anmeldung, für die Gebühren entrichtet worden sind, nämlich auf die Ansprüche Nr.
4. ☐ Der Anmelder hat die erforderlichen zusätzlichen Recherchegebühren nicht rechtzeitig entrichtet. Der internationale Recherchenbericht beschränkt sich daher auf die in den Ansprüchen zuerst erwähnte Erfindung; diese ist in folgenden Ansprüchen erfaßt:

Bemerkungen hinsichtlich eines Widerspruchs

- ☐ Die zusätzlichen Gebühren wurden vom Anmelder unter Widerspruch gezahlt.
- ☒ Die Zahlung zusätzlicher Gebühren erfolgte ohne Widerspruch.

WEITERE ANGABEN

PCT/ISA/ 210

1. Ansprüche: 1-35, teilweise 37-40

DNA-Sequenz das gp190/MSP-1 Oberflächenprotein von Plasmodium kodierend, Wirtsorganismus diese Sequenz enthaltend, Verwendung des gp190/MSP-1 Oberflächenproteins oder der dafür kodierenden DNA zur Immunisierung gegen Malaria, Vektor besagte DNA-Sequenz enthaltend, Impfstoff das gp190/MSP-1 Oberflächenprotein oder die dafür kodierende DNA enthaltend, Verfahren zur Herstellung des gp190/MSP-1 Oberflächenproteins

2. Ansprüche: 36, teilweise 37-40

Verfahren zur Stabilisierung von Gen-Sequenzen durch Verringerung des AT-Gehaltes der Sequenzen, stabilisiertes Gen, Vektor dieses Gen enthaltend, Impfstoff diesen Vektor enthaltend

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP 97/05441

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9428930 A	22-12-94	AU 7060294 A EP 0717636 A	03-01-95 26-06-96
EP 0340359 A	08-11-89	AU 620041 B AU 2156988 A CA 1331155 A DE 3882522 A DE 3882522 T DK 478588 A ES 2058294 T IE 62490 B JP 2115464 C JP 2167088 A JP 8013275 B PT 88362 A,B US 5147788 A	13-02-92 09-11-89 02-08-94 26-08-93 03-03-94 07-11-89 01-11-94 08-02-95 06-12-96 27-06-90 14-02-96 30-11-89 15-09-92
EP 0154454 A	11-09-85	AU 592749 B AU 3904685 A DK 79985 A GB 2154592 A IL 74409 A JP 2584733 B JP 61019490 A JP 6189772 A PH 25993 A US 5597708 A	25-01-90 05-09-85 23-08-85 11-09-85 26-08-94 26-02-97 28-01-86 12-07-94 13-01-92 28-01-97
EP 0385962 A	05-09-90	AU 638438 B AU 5163090 A CA 2024811 A EP 0413019 A IL 93513 A JP 3504333 T WO 9010076 A US 5500365 A	01-07-93 26-09-90 25-08-90 20-02-91 14-11-95 26-09-91 07-09-90 19-03-96
EP 0359472 A	21-03-90	AT 132193 T AU 623429 B AU 4118289 A	15-01-96 14-05-92 15-03-90

INTERNATIONALER RECHTENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Intern **laufs Aktenzeichen**

PCT/EP 97/05441

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
EP 0359472 A		CN 1044298 A	01-08-90
		DE 68925253 D	08-02-96
		EP 0682115 A	15-11-95
		ES 2083384 T	16-04-96
		JP 2186989 A	23-07-90
		US 5380831 A	10-01-95
		US 5567600 A	22-10-96
		US 5567862 A	22-10-96

THIS PAGE BLANK (USPTO)

PCTWELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales BüroINTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation ⁶: C12N 15/30, C07K 14/445, C12N 15/62, A61K 39/015, 31/70, C07H 21/00, C12N 1/21, 5/10, 1/11, 15/67	A2	(11) Internationale Veröffentlichungsnummer: WO 98/14583 (43) Internationales Veröffentlichungsdatum: 9. April 1998 (09.04.98)
(21) Internationales Aktenzeichen: PCT/EP97/05441 (22) Internationales Anmeldedatum: 2. Oktober 1997 (02.10.97) (30) Prioritätsdaten: 196 40 817.2 2. Oktober 1996 (02.10.96) DE (71)(72) Anmelder und Erfinder: BUJARD, Hermann [DE/DE]; Remlerstrasse 9, D-69120 Heidelberg (DE). (72) Erfinder; und (75) Erfinder/Anmelder (nur für US): TOLLE, Ralf [DE/DE]; Friedrich-Naumann-Strasse 8, D-71636 Ludwigsburg (DE). PAN, Weiqing [CN/DE]; Im Buschgewann 71, D-69123 Heidelberg (DE). (74) Anwalt: GRÜNECKER, KINKELDEY, STOCKMAIR & SCHWANHÄUSSER; Maximilianstrasse 58, D-80538 München (DE).		(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Veröffentlicht <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i>
(54) Title: METHOD FOR PRODUCING RECOMBINANTS INTENDED FOR USE IN A COMPLETE MALARIA ANTIGENE GP190/MSP1		
(54) Bezeichnung: REKOMBINANTES HERSTELLUNGSVERFAHREN FÜR EIN VOLLSTÄNDIGES MALARIA-ANTIGEN GP190/MSP1		
(57) Abstract		
<p>The present invention relates to a method for producing recombinants intended for use in the complete cell-surface protein gp190/MSP1 from plasmodium, especially plasmodium falciparum, as well as the complete DNA sequence of this protein and the appropriate host organisms suited for expressing said sequence, whereby the protein concerned can be entirely synthesized outside the parasite. Also, the inventive method enables sufficient production of above-mentioned protein and its supply as a vaccine. Finally disclosed is a process for stabilizing genes with high At concentration.</p>		
(57) Zusammenfassung		
<p>Die Erfindung betrifft ein rekombinantes Herstellungsverfahren für das komplette gp190/MSP1-Oberflächenprotein von Plasmodium, insbesondere Plasmodium falciparum, sowie die vollständige DNA-Sequenz dieses Proteins und geeignete Wirtsorganismen für die Expression der Sequenz, wodurch das Protein in seiner Gesamtheit erstmals außerhalb des Parasiten synthetisiert werden konnte. Die Erfindung eröffnet erstmals die Möglichkeit, das gp190/MSP1-Oberflächenprotein in ausreichender Menge herzustellen; ferner ist es ein Gegenstand der Erfindung gp190/MSP1 als Impfstoff zur Verfügung zu stellen. Schließlich gibt die Erfindung ein Verfahren zur Stabilisierung von AT-reichen Genen an.</p>		

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidshan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland			TR	Türkei
BG	Bulgarien	HU	Ungarn	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
BY	Belarus	IS	Island	MW	Malawi	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	MX	Mexiko	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	VN	Vietnam
CG	Kongo	KE	Kenia	NL	Niederlande	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	NZ	Neuseeland		
CM	Kamerun			PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	KZ	Kasachstan	RO	Rumänien		
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		

Rekombinantes Herstellungsverfahren für ein vollständiges Malaria-Antigen gp190/MSP1

Die Erfindung betrifft ein rekombinantes Herstellungsverfahren für das vollständige Malaria-Antigen gp190/MSP1 sowie einzelner natürlich vorkommender Domänen und Teile derselben durch Expression (einer) synthetischer DNA-Sequenzen. Die Erfindung betrifft außerdem die durch das Verfahren hergestellten DNA-Sequenzen und die für die Expression der DNA-Sequenzen verwendeten Wirtsorganismen. Ferner betrifft die Erfindung die Verwendung des vollständigen Malaria-Antigens sowie Teile derselben als Impfstoff zur Immunisierung gegen Malaria.

Schließlich betrifft die vorliegende Erfindung ein Stabilisierungsverfahren für AT-reiche Gene, sowie stabilisierte Gene, die sich durch einen geringeren AT-Gehalt auszeichnen.

Malaria ist weltweit eine der bedeutendsten Infektionskrankheiten. Nach Angaben der WHO waren 1990 in 99 Ländern 40% der Weltbevölkerung einem Malariarisiko ausgesetzt. Ihre Verbreitung nimmt derzeit wieder massiv zu. Dies ist vor allem auf eine intensive Resistenzbildung der Malariaerreger zurückzuführen, die dadurch gefördert wird, daß die zur Therapie eingesetzten Medikamente ebenfalls als Prophylaxe empfohlen und eingenommen werden. Neben der Suche nach neuen, wirksamen Chemotherapeutika werden heute Hoffnungen auch in die Entwicklung von Impfstoffen gesetzt, da Menschen in Malaria-epidemischen Regionen der Welt verschiedene Arten von Immunität zu entwickeln vermögen. Neben einer natürlichen Resistenz gegen Malaria, die sich bei heterozygoten Trägern des Sichelzellgens und bei Personen mit Thalassämie und Glucose-6-Phosphat-Dehydrogenase-Mangel ausbildet, können im Laufe einer Malariainfektion im Menschen Immunitätsmechanismen einsetzen, die sich in einer erhöhten Widerstandsfähigkeit gegenüber den Plasmodien äußern. Dementsprechend ist der Krankheitsverlauf in stark durchseuchten Bevölkerungspopulationen weitaus weniger bedrohlich als bei Personen, die der Infektion weniger häufig oder erstmalig ausgesetzt sind.

Das Hauptproblem bei der Impfstoffentwicklung ist die Identifikation eines Antigens, das schützende Immunität bewirken kann, da kein leicht zugängliches und gut defi-

niertes Tiermodell für die vier den Menschen befallenden Parasiten vorhanden ist. Die Malaria-Erreger gehören der Gruppe der Plasmodien an, wobei die Infektion mit einem der vier Erreger *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* oder *Plasmodium falciparum* durch den Stich von Anopheles-Mücken erfolgt. Von diesem Erreger ist *Plasmodium falciparum* der gefährlichste und der am weitesten verbreitete.

Das Hauptoberflächenprotein von Merozoiten, der invasiven Form der Blutstadien des Malaria-Erregers *Plasmodium falciparum* und anderer Malaria-Erreger wie *P. vivax*, ist ein 190 - 220 kD Glykoprotein. Spät in der Entwicklung des Parasiten wird dieser Vorläufer in kleinere Proteine prozessiert, die jedoch als einheitlicher Komplex aus Merozoiten isoliert werden können. Der Komplex ist mittels eines Glycosylphosphatidyl-Inositol-Ankers mit der Merozoitenmembran verknüpft. Die Sequenzen der gp 190-Proteine verschiedener *P. falciparum*-Stämme fallen in zwei Gruppen, zwischen denen intragene Rekombination häufig ist. Insgesamt besteht das Protein aus mehreren hochkonservierten Regionen, aus einem dimorphen Bereich, welche jeweils einem von zwei Allelen angehören und aus zwei relativ kleinen oligomorphen Blöcken im N-terminalen Bereich (Tanabe, K., Mackay, M., Goman, M. und Scaife, J.G. (1987), Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. J. Mol. Biol. 195, 273-287; Miller, L.H., Roberts, T., Shaha-buddin, M. und McCutchan, T.F. (1993), Analysis of sequence diversity in the *Plasmodium falciparum* merozoite surface protein-1 (MSP-1). Mol. Biochem. Parasitol. 59, 1-14).

Das gp190/MSP1 galt bereits früh als ein möglicher Kandidat für einen Impfstoff. So wurde im Nagermodell nach Immunisierung mit dem analogen Protein aktiver Schutz gegen Infektion mit Nager-Parasiten erhalten. Passiver Schutz ließ sich mit gegen dieses Protein gerichteten Antikörpern erreichen (siehe auch Holder, A. A. und Freeman, R.R. (1981), Immunization against blood-stage rodent malaria using purified parasite antigens, Nature 294, 361-364; Majarian, W.R., Daly, T. M., Weidanz, W.P. und Long, C.A. (1984), Passive immunization against murine malaria with an IgG3 monoclonal antibody, J. Immunol. 132, 3131-3137). Die Daten, die diese Annahme belegen sollen, sind im einzelnen jedoch statistisch nicht signifikant.

Darüber hinaus sind mehrere monoklonale Antikörper, welche in vitro die Invasion von Erythrozyten durch *P. falciparum* inhibieren, sind gegen gp190/MSP1 gerichtet (Pirson, P.J. und Perkins, M.E. (1985), Characterization with monoclonal antibodies of a surface antigen of *Plasmodium falciparum* merozoites. J. Immunol. 134, 1946-1951; Blackman, M.J., Heidrich, H.-G., Donachie, S., McBride, J.S. und Holder A.A. (1990), A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies, J. Exp. Med. 172, 379-382).

Schließlich wurde eine Reihe von Impfstudien mit gp190/MSP1-Material aus *P. falciparum* an Primaten, insbesondere an Aotus und Saimiri-Affen durchgeführt (siehe auch Perrin, L.H., Merkli, B., Loche, M., Chizzolini, C., Smart, J. und Richle, R. (1984), Antimalarial immunity in Saimiri monkeys. Immunization with surface components of asexual blood stages, J. Exp. Med. 160, 441-451; Hall, R., Hyde, J.E., Goman, M., Simmons, D.L., Hope, I.A., Mackay, M. und Scaife, J.G. (1984), Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria, Nature 311, 379-382; Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chari, E.B.T. und Kan, S.-C. (1987), Merozoite surface coat precursor protein completely protects Aotus monkeys against *Plasmodium falciparum* malaria, Proc. Natl. Acad. Sci. USA 84, 3014-3018; Ettlinger, H.M., Caspers, P., Materile, H., Schoenfeld, H.-J., Stueber, D. und Takacs, B. (1991), Ability of recombinant or native proteins to protect monkeys against heterologous challenge with *Plasmodium falciparum*, Inf. Imm. 59, 3498-3503; Holder, A.A., Freeman, R.R. und Nicholls, S.C. (1988), Immunization against *Plasmodium falciparum* with recombinant polypeptides produced in *Escherichia coli*, Parasite Immunol. 10, 607-617; Herrera, S., Herrera, M.A., Perlaza, B.L., Burki, Y., Caspers, P., Döbeli, H., Rotmann, D. und Certa, U. (1990), Immunization of Aotus monkeys with *Plasmodium falciparum* blood-stage recombinant proteins, Proc. Natl. Acad. Sci. USA 87, 4017-4021; Herrera, M.A., Rosero, F., Herrera, S., Caspers, P., Rotmann, D., Sinigaglia, F. und Certa, U. (1992), Protection against malaria in Aotus monkeys immunized with a recombinant blood-stage antigen fused to a universal T-cell epitope; correlation of serum gamma interferon levels with protection, Inf. Imm. 60, 154-158; Patarroyo, M.E., Ro-

mero P., Torres, M.L., Clavijo, P., Moreno, A., Martinez, A., Rodriguez, R., Guzmán, F. und Cabezas, E. (1987), Induction of protective immunity against experimental infection with malaria using synthetic peptides, Nature 328, 629-632). In diesen Impfstudien lassen sich hierbei zwei Ansätze unterscheiden

- Verwendung von aus Parasiten isoliertem Material und
- Einsatz von in heterologen Expressionssystemen gewonnenem Material.

Letzteres bestand in der Regel aus relativ kleinen Teilbereichen des Gesamtproteins. Obwohl die ersten Resultate der in Vorversuchen an Affen durchgeführten Impfungen andeuten, daß gp190/MSP1 einen Schutz vermitteln könnte, haben alle an Primaten durchgeführten Experimente zwei Probleme, welche einen solchen Schluß in Frage stellen:

- (a) sie wurden an zu kleinen Tiergruppen durchgeführt
- (b) sie wurden nicht wiederholt.

Die Resultate und die daraus gezogenen Schlüsse sind daher statistisch nicht abgesichert. Neben dem schwierigen Zugang zu geeigneten Affen liegt das zugrunde liegende Hauptproblem darin, daß es bislang nicht möglich war, gutes Impfmateri al in ausreichender Menge herzustellen.

Andererseits konnten nach der Sequenzierung des gp190-Gens aus dem K1- und dem MAD20-Stamm von *Plasmodium falciparum* konnten überlappende Fragmente in *E. coli* exprimiert werden. Mit diesem Material zeigten epidemiologische Studien in Westafrika, daß in der Gruppe der Adoleszenten eine Korrelation bestand zwischen Antikörpertiter gegen gp190/MSP1-Fragmente einerseits und einem Schutz vor Parasiteninfektion andererseits. Darüber hinaus schien der Titer auch mit der Fähigkeit zu korrelieren, die Parasitämie auch auf niedrigem Niveau zu kontrollieren (Tolle et al. (1993): A prospective study of the association between the human humoral immune response to *Plasmodium falciparum* blood stage antigen gp190 and control of malarial infections. Infect Immun. 61, 40-47). Diese Resultate werden ergänzt durch neue Untersuchungen an Aotusaffen im Rahmen der vorliegenden Erfindung. Hier wurde ein hoher Schutz gegen Infektion mit dem Parasiten dadurch erreicht, daß Proteinpräparationen aus *Plasmodium falciparum*, die überwiegend aus nichtprozessiertem

gp190/MSP1 bestanden, als Impfstoff benutzt worden waren. Die Affen mit dem höchsten Antikörpertiter gegen gp190/MSP1 waren am besten geschützt. Diese Resultate machten letztendlich das gp190 zu einem vielversprechenden Kandidaten für eine Impfstoff gegen *Malaria tropica*.

Von einigen Arbeitsgruppen wurde der C-terminalen Domäne des gp190 (p19 bzw. p42) eine besondere Rolle bei der gp190 vermittelten Immunität zugewiesen (siehe auch Chang, S.P., Case, S.E., Gosnell, W.L., Hashimoto, A., Kramer, K.J., Tam, L. Q., Hashiro, C.Q., Nikaido, C.M., Gibson, H.L., Lee-Ng, C.T., Barr, P.J., Yokota, B.T. und Hui, G.S.N. (1996), A recombinant baculovirus 42-kilodalton C-terminal fragment of *Plasmodium falciparum* merozoite surface protein 1 protects Aotus monkeys against malaria, *Inf. Imm.* 64, 253-261; Burghaus, P.A., Welde, B.T., Hall, T., Richards, R.L., Egan, A.F., Riley, E.M., Ripley-Ballou, W. und Holder A.A. (1996), Immunization of Aotus nancymai with recombinant C-terminus of *Plasmodium falciparum* merozoite surface protein-1 in liposomes and alum adjuvant does not induce protection against a challenge infection, *Inf. Imm.*, in press).

Bislang ist es jedoch nicht möglich, auf rationaler Basis andere Teile des gp190 als nicht relevant für eine schützende Immunantwort auszuschließen. Es ist daher nach wie vor notwendig, das gesamte Gen bzw. das intakte gp190 für Impfversuche zu verwenden. Trotz mehrfacher Versuche verschiedener Arbeitskreise ist es jedoch noch nicht gelungen, das gesamte Gen des gp190/MSP1 zu klonieren und zu exprimieren.

Bislang war es auch noch nicht möglich, a priori einen Teil aus der gp190-Sequenz für die schützende Immunantwort als nicht relevant auszuschließen, so daß es nach wie vor notwendig ist, das gesamte Gen bzw. das gesamte Genprodukt für Impfversuche zu verwenden. Es ist jedoch trotz vieler Versuche mehrerer Arbeitskreise noch nicht gelungen, das gesamte Gen des gp190/MSP1 zu klonieren.

Es war daher eine Aufgabe der vorliegenden Erfindung, Impfmateriel in Form von vollständigem gp190/MSP1 in ausreichender Menge zur Verfügung zu stellen. Es war

eine weitere Aufgabe der vorliegenden Erfindung, ein Verfahren anzugeben, mit dem dieses Impfmateriel gewonnen werden kann.

Es war außerdem eine weitere Aufgabe gemäß der vorliegenden Erfindung, eine vollständige DNA-Sequenz von gp190/MSP1 anzugeben, die in einem Wirtsorganismus exprimierbar ist.

Weiterhin war es eine Aufgabe der vorliegenden Erfindung, Wirtsorganismen anzugeben, die das vollständige Gen gp190/MSP1 enthalten.

Schließlich war es auch eine Aufgabe der vorliegenden Erfindung, ein Stabilisierungsverfahren für AT-reiche Gene anzugeben, sowie ein für die Expression geeignetes, stabilisiertes Gen, das sich durch eine Erniedrigung des AT-Gehalts auszeichnet.

Diese Aufgaben werden durch die in den Ansprüchen angegebenen Gegenstände gelöst.

Im folgenden werden einige Begriffe näher erläutert, um klarzustellen, wie sie in diesem Zusammenhang verstanden werden sollen.

"Rekombinantes Herstellungsverfahren" bedeutet, daß ein Protein von einer DNA-Sequenz durch einen geeigneten Wirtsorganismus exprimiert wird, wobei die DNA-Sequenz aus einer Klonierung und Fusion einzelner DNA-Abschnitte entstanden ist.

"Vollständiges gp190/MSP1-Protein" meint hier das gesamte, aus o.g. Plasmodien, insbesondere *Plasmodium falciparum*, isolierbare gp190/MSP1-Oberflächenprotein, das das Hauptoberflächenprotein der Merozoiten des o.g. Erregers darstellt sowie die Proteine mit analoger Funktion aus den anderen Plasmodiumarten, wie *P. vivax*. Der Begriff betrifft somit jeweils das Hauptoberflächenprotein der Merozoiten der 4 o.g. für den Menschen gefährlichen Malariaerreger. "Vollständiges gp190/MSP1-Gen" ist das für dieses Protein kodierende Gen. "Vollständig" bedeutet in diesem Zusammenhang, daß die gesamte Aminosäuresequenz des nativen Proteins vorhanden ist, bzw. daß die Gensequenz für die gesamte Aminosäuresequenz des nativen Proteins kodiert.

Mit umfaßt sind jedoch auch mutierte und/oder verkürzte Formen von gp190/MSP1, sofern sie das gleiche Immunisierungspotential (Impfschutz) wie das vollständige gp190/MSP1 aufweisen. Schließlich umfaßt der Begriff auch Varianten von gp190/MSP1 die sich dadurch auszeichnen, daß sie Proteinabschnitte verschiedener Allele in einem Proteinmolekül enthalten.

"FCB-1" ist ein Stamm von *P. falciparum*, wie beschrieben bei Heidrich, H.-G., Miettinen-Baumann, A., Eckerskorn, C. und Lottspeich, F. (1989) The N-terminal amino acid sequences of the Plasmodium falciparum (FCB1) merozoite surface antigens of 42 and 36 kilodalton, both derived from the 185-195-kilodalton precursor. Mol. Biochem. Parasitol. 34, 147-154.

"Ankersignal" meint hier eine Proteinstruktur, für die eine DNA-Sequenz am 3'- oder 5'-Ende des erfindungsgemäßen Gens kodiert. Ankersignale sind Strukturen, die einem Polypeptid die Verankerung an anderen Strukturen, wie z.B. Membranen ermöglichen.

"Signalpeptid" bedeutet hier eine Proteinstruktur, für die eine DNA-Sequenz am N-terminalen Ende des erfindungsgemäßen Gens kodiert. Signalpeptide sind Strukturen, die dem Polypeptid u.a. ein Einschleusen in Membranen ermöglichen.

"AT-Gehalt" meint im Zusammenhang der vorliegenden Erfindung die prozentuale Menge von Adenin/Thymin-Basenpaaren im Verhältnis zu Guanin/Cytosin-Basenpaaren.

"Klonierung" soll hier alle im Stand der Technik bekannten Klonierungsmethoden umfassen, die hier zum Einsatz kommen könnten, die jedoch nicht alle im einzelnen beschrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

"Expression in einem geeigneten Expressionssystem" soll hier alle im Stand der Technik bekannten Expressionsmethoden in bekannten Expressionssystemen umfassen, die hier zum Einsatz kommen könnten, die jedoch nicht alle im einzelnen be-

schrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

Es ist eine erste Aufgabe gemäß der vorliegenden Erfindung, ein Verfahren anzugeben, durch das das gp190/MSP1-Protein und das Gen hierfür in ausreichender Menge ohne übermäßige Kosten produziert werden kann.

Diese Aufgabe wird durch das in Anspruch 1 beschriebene rekombinante Herstellungsverfahren gelöst, durch das ein voll-ständiges gp190/MSP-1-Gen und das davon kodierte Protein in ausreichender Menge erhältlich ist.

Durch dieses Verfahren wurde es erstmals möglich, das Protein in seiner Gesamtheit außerhalb des Parasiten zu synthetisieren. Das so synthetisierte Protein ist, wie die Analyse mit konformationelle Epitope erkennenden monoklonalen Antikörpern zeigt, zumindest über weite Bereiche in natürlich gefalteter Form herstellbar. Durch das rekombinante Herstellungsverfahren konnten jeweils mehrere Milligramm intaktes gp190/MSP1 aus den Wirtsorganismen gewonnen werden, eine Menge, die aus technischen und aus Kostengründen aus Parasiten praktisch nicht gewonnen werden kann. Die jetzt mögliche Produktion des Proteins in beliebigen Mengen eröffnet neue Perspektiven für seinen Einsatz als experimentellen Impfstoff gegen Malaria. Darüber hinaus ist der Weg frei für die Entwicklung von Lebendimpfstoffen sowie für Vakzine auf Nukleinsäure-Basis.

Vorzugsweise liegt der Synthese der für das Protein gp190/MSP1 kodierenden Gen-Sequenz die Sequenz des *P. falciparum* FCB-1 Stammes zugrunde. *P. falciparum* ist der Erreger der Malaria tropica und damit der gefährlichste unter den Malaria-Arten. Das zugrunde liegende Gen ist ein Vertreter des "K1-Allels", wobei K1 für einen bestimmten *P. falciparum*-Stamm steht. Seine kodierende Sequenz erstreckt sich über 4917 Basenpaare und schließt eine Signalsequenz am N-terminalen Ende sowie eine Anker-Sequenz am C-terminalen Ende ein.

Weiterhin ist das rekombinante Herstellungsverfahren gemäß der Erfindung vorzugsweise dadurch gekennzeichnet, daß der AT-Gehalt der dem Protein zugrunde

liegenden DNA-Sequenz gegenüber der Wildtyp-Sequenz erniedrigt ist, von 74% im ursprünglichen Gen auf vorzugsweise ca. 55%, indem bspw. unter Erhalt der Aminosäuresequenz des FCB-1-Proteins eine DNA-Sequenz mit den im menschlichen Genom üblichen Codon-Häufigkeiten hergestellt wird. Andere Codonhäufigkeiten, welche den AT-Gehalt erniedrigen, sind ebenfalls denkbar.

Vorzugsweise kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz einschließlich Signalpeptid und GPI-Ankersignalpeptid, im weiteren als gp190^s bezeichnet.

In einer weiteren bevorzugten Ausführungsform kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des GPI-Ankersignals. Diese Ausführungsform wird im weiteren als gp190^{s1} bezeichnet.

In noch einer weiteren bevorzugten Ausführung kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des GPI-Ankersignals und des Signalpeptids. Diese Ausführungsform wird im weiteren als gp190^{s2} bezeichnet.

In einer weiteren bevorzugten Ausführungsform kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz und eine Transmembranankersequenz.

In einer besonders bevorzugten Ausführungsform umfaßt das rekombinante Herstellungsverfahren folgende Schritte.

Zunächst den Entwurf der zu synthetisierenden DNA-Sequenz auf der Basis des Gens aus *P. falciparum* FCB-1, wobei eine DNA-Sequenz mit den z.B. im menschlichen Genom üblichen Codon-Häufigkeiten unter Beibehalten der Aminosäuresequenz des FCB-1-Proteins hergestellt wird.

Hierdurch sollte der AT-Gehalt des Gens reduziert werden, vorzugsweise auf 55%.

Im weiteren Verfahren wird die entworfene Sequenz bspw. in 5 überlappende Regionen eingeteilt, welche jeweils Domänen der natürlichen Prozessierungsprodukte des gp190/MSP1-Proteins aus FCB-1 entsprechen: p83, p31, p36, p30 und p19.

Es werden Desoxyoligonukleotide synthetisiert, die jeweils mindestens die gesamte Länge einer Region abdecken.

Besonders bevorzugt werden die Desoxyoligonukleotide so synthetisiert, daß ihre Sequenz abwechselnd dem "oberen" (5' -3') bzw. dem "unteren" (3' - 5') DNA-Strang entspricht. Die Länge dieser Oligonukleotide ist vorzugsweise durchschnittlich 120 Nukleotide und sie überlappt die benachbarten Sequenzen jeweils um ca. 20 Basen.

In einer möglichen Ausführungsform werden DNA-Sequenzen von etwa doppelter Länge wie die jeweiligen Ausgangsprodukte durch asymmetrische PCR hergestellt und zwar so, daß die überschüssigen DNA-Sequenzen, die benachbart sind, jeweils den gegenüberliegenden Strang repräsentieren. Dies führt in einem zweiten

PCR-Amplifikationszyklus zu einem Zweitprodukt, das der Länge von vier ursprünglich eingesetzten Oligonukleotiden (abzüglich der überlappenden Region) entspricht. Die Überführung dieser Produkte in ein überwiegend aus Einzelstrang-DNA bestehendes Präparat durch asymmetrische PCR mit den endständigen Oligonukleotiden erlaubt in einem weiteren Amplifikationsschritt die Herstellung eines 800 bp langen doppelsträngigen DNA-Fragments in nur 25 PCR-Zyklen.

Auf diese Weise werden direkt die kodierenden Regionen für p19, p30, p36 und p31 synthetisiert und in *E. coli* molekular kloniert. Klone mit fehlerfreier Sequenz wurden entweder direkt oder durch Zusammensetzen fehlerfreier Teilsequenzen erhalten. Die Region, welche das p83 kodiert, wurde durch Fusion aus zwei etwa 1200 bp umfassenden Sequenzen erhalten.

Im weiteren Verlauf des Verfahrens wurden die einzelnen Sequenzen kloniert. Als Expressionsvektoren bieten sich vorzugsweise die Plasmide pDS56, RBSII ("Hochuli,

E., Bannwarth, W. Döbeli, H. Gentz, R., and Stüber, D. (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Biotechn. 6*, 1321-1325"), pBi-5 ("Baron, U., Freundlieb, S., Gossen, M. and Bujard, H. (1995) Corregulation of two gene activities by tetracycline via a bidirectional promoter. *Nucl. Acids Res. 23*, 3605-3606") und ppTMCS an. Es sind jedoch auch andere Expressionsvektoren denkbar.

Bevorzugte Wirtsorganismen für die Expression sind *E. coli*, besonders bevorzugt der Stamm DH5alphaZ1 (R. Rutz, Dissertation 1996, Universität Heidelberg), HeLa-Zellen, CHO-Zellen, *Toxoplasma gondii* (Pfefferkorn, E.R. and Pfefferkorn, C.C. 1976, *Toxoplasma gondii*: Isolation and preliminary characterization of temperature - sensitive mutants. *Exp. Parasitol. 39*, 365-376) oder *Leishmania*. Weitere Wirtssysteme könnten z.B. Hefen, Baculoviren oder Adenoviren sein, wobei der Gegenstand der Erfindung nicht auf die genannten Wirtssysteme beschränkt sein sollte.

Es war eine weitere Aufgabe der vorliegenden Erfindung, eine vollständige, zur Expression geeignete DNA-Sequenz des gp190/MSP1-Oberflächenproteins von *P. falciparum* anzugeben.

Diese Aufgabe wird durch die in Anspruch 17 genannte Erfindung gelöst, wobei die Sequenz durch das oben beschriebene rekombinante Herstellungsverfahren erhältlich ist.

In einer bevorzugten Ausführungsform der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz.

In einer anderen bevorzugten Ausführungsform der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz mit Ausnahme des Ankersignals.

In einer weiteren bevorzugten Ausführungsform gemäß der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz mit Ausnahme des Ankersignals und des Signalpeptids. Diese Ausführungs-

form von gp190/MSP1 kann dadurch gekennzeichnet sein, daß sie am N-Terminus 11 zusätzliche Aminosäuren, davon 6 Histidine, enthält.

Besonders bevorzugt enthält die zur Expression geeignete DNA-Sequenz keine erkennbaren "splice-donor" und "splice-acceptor"-Signale, und sie ist vorzugsweise dadurch gekennzeichnet, daß sie keine größeren GC-reichen Sequenzen enthält, die stabile Haarnadelstrukturen auf RNA-Ebene bewirken könnten.

Vorzugsweise sollten Erkennungssignale für Restriktionsenzyme, welche Sequenzen von sechs und mehr Basenpaaren erkennen, vermieden werden.

In einer bevorzugten Ausführung werden spezifische, d.h. nur einmal im Gen vorkommende Schnittstellen für Restriktionsendonukleasen in Regionen eingeführt, welche die nach der Prozessierung des Proteins entstehenden Domänen trennen.

Besonders bevorzugt sollten an beiden Enden des Gens Sequenzen für Restriktionsendonukleasen vorhanden sein, die im Gen nicht vorkommen.

Weiterhin werden durch die Erfindung Wirtsorganismen zur Verfügung gestellt, die die vollständige Sequenz des gp190/MSP-1 Oberflächenproteins enthalten.

Solche Wirtsorganismen sind vorzugsweise *E. coli*, besonders bevorzugt der Stamm DH5alphaZ1, HeLa-Zellen, CHO-Zellen, *Toxoplasma gondii* oder *Leishmania*. Die HeLa-Zellen und CHO-Zellen sollten vorzugsweise konstitutiv tTA synthetisieren.

Schließlich stellt die vorliegende Erfindung eine Möglichkeit zur Verfügung, ein nach dem rekombinanten Herstellungsverfahren erzeugtes gp190/MSP1-Oberflächenprotein oder Teile desselben zur aktiven Immunisierung gegen Malaria zu verwenden.

Das hier vorgestellte Syntheschema erlaubt auch das zweite Allel des gp190/MSP1-Gens herzustellen. Damit wird dem Dimorphismus des Proteins Rechnung getragen. Die Hauptvariabilität des Proteins beruht jedoch auf den Sequenzen von zwei relativ kurzen Blöcken (Block II und IV, Ref. 1), die oligomorph sind. Die vor-

liegenden Sequenzdaten ermöglichen es, daß mit 6-8 Sequenzkombinationen dieser Blöcke über 95% aller bekannten gp190/MSP1-Sequenzen abgedeckt werden können. Die Synthese dieser Sequenzvarianten läßt sich problemlos anhand der hier vorgestellten Strategien verwirklichen, so daß Varianten sowohl in das K1 als auch in das MAD20-Allel eingebaut werden können. Impfstoffe aus den hierdurch entstehenden Sequenzfamilien können ggf. gegen ein breites Spektrum von Parasiten mit gp190/MSP1-Varianten Schutz verleihen.

Die Herstellung von verschiedenen Impfstofftypen ist möglich:

- Auf der Ebene von Proteinpräparaten, wobei jeweils Mischungen der zwei Familien (K1-Typ, MAD20-Typ mit verschiedenen Varianten der Blöcke II und IV) zur Anwendung kommen können. Verschiedene Träger bzw. Adjuvans-Materialien können zum Einsatz kommen: Aluminiumoxid, Liposomen, Iscoms QSz1 etc.
- Auf der Ebene der Lebendimpfstoffe: (a) virale Träger, insbesondere Vakzinia und Adenoviren; (b) Parasiten als Träger, insbesondere avirulente Formen von Leishmania und Toxoplasma; (c) bakterielle Träger, z.B. Salmonella
- Auf der Ebene der Nukleinsäuren, wobei bspw. für die Gentherapie geeignete Vektoren als Vehikel zum Einbringen der Gene in den Wirt verwendet werden; weiterhin ist das Einbringen von Ribonukleinsäuren, welche das gewünschte Protein kodieren, denkbar.

Eine weitere Möglichkeit der Impfung besteht in der Verwendung eines gemäß dem erfindungsgemäßen rekombinanten Herstellungsverfahren erzeugten gp190/MSP1-Proteins zur Herstellung von monoklonalen Antikörpern, die dann ihrerseits zur passiven Immunisierung gegen Malaria verwendet werden.

Ebenso wird eine Verwendung der gemäß dem rekombinanten Herstellungsverfahren in einem Zwischenschritt entstandenen, dem Protein zugrunde liegenden DNA-Sequenz zur Herstellung einer Vakzine auf Nukleinsäurebasis ermöglicht.

Schließlich betrifft die Erfindung auch ein Verfahren zur Stabilisierung von Gen-Sequenzen, insbesondere von Sequenzen, die keine ausreichende Stabilität in Expressionssystemen zeigen.

Diese Stabilisierung wird erfindungsgemäß dadurch erreicht, daß der AT-Gehalt der Sequenz verringert wird.

Weiterhin wird durch die Erfindung ein stabilisiertes Gen zur Verfügung gestellt, das sich dadurch auszeichnet, daß es einen geringeren AT-Gehalt aufweist. Ein Beispiel für ein solches, stabilisiertes Gen ist das Gen für das gp190/MSP1-Oberflächenprotein gemäß der vorliegenden Erfindung.

Im folgenden soll die Erfindung anhand der Abbildungen und Tabellen sowie einiger Beispiele in einzelnen Ausführungsformen beschrieben werden.

Dabei zeigt:

Abb. 1: Schematische Darstellung des gp190/MSP1 Vorläuferproteins aus *P. falciparum* (FCB-1).

Abb. 2: Zwei mit nativem gp190/MSP1 aus *P. falciparum* (FCB-1) an Aotus-Affen durchgeführte Impfversuche.

Abb. 2A: mit 3 x 60 Mikrogramm gp190/MSP1

Abb. 2B: mit 3 x 40 Mikrogramm gp190/MSP1

Abb. 3A: Strategie der Synthese des gp190/MSP1-Gens

Abb. 3B: Prinzip der PCR-gestützten Total-Synthese

Abb. 3C: Totalsequenz des gp190^S

Abb. 3D: N- und C-Terminus der gp190^{S1}-Variante

Abb. 4A: Expressionsvektor pDS56 mit gp190^{S2}-Sequenz

Abb. 4B: Gelelektrophorese von gp190^{S2}.

Abb. 5A: Expressionsvektor pBi-5 mit gp190^{S1}-Sequenz

Abb. 5B: Immunfluoreszenz von HeLa-Zellen

Abb. 5C: Elektrophoretische Charakterisierung von aus HeLa-Zellen aufgereinigtem gp190^{S1}

Abb. 6A: Expressionsvektor ppT 190 mit gp190-Sequenz

Abb. 6B: Immunfluoreszenz der Expression von gp190^S in *T. gondii*

Abb. 6C: Polyacrylamid-Gelelektrophorese von gp190 aus *T. gondii*

Bei dem in Abb. 1 schematisch dargestellten gp190/MSP1-Vorläuferprotein aus *P. falciparum* (FCB-1) repräsentieren die dunklen Blöcke Regionen, die in allen Stämmen hochkonserviert vorliegen. Die schraffierten Blöcke zeigen die dimorphen Bereiche, welche im Falle des FCB-1 Isolates dem K1-Allel entstammen. O1 und O2 zeigen die oligomorphen Bereiche. S bezeichnet die Signalpeptidsequenz, welche 19 Aminosäuren umfaßt, GA, die C-terminale Region, welche das Signal für die GPI-Verankerung des Proteins in der Membran enthält. Die Pfeile deuten die Prozessierungsstellen an, durch die Proteine p53, p31, p36, p30 und p19 entstehen. Das gp190-Gen kodiert insgesamt 1639 Aminosäuren.

Die weiteren Abbildungen werden im Zusammenhang mit den folgenden Beispielen ausführlicher erläutert.

BEISPIELE

Beispiel 1: Totalsynthese einer das gp190/MSP1 kodierenden DNA-Sequenz (siehe hierzu Abb. 3)

A. Strategie der Synthese des gp190/MSP1-Gens (gp190^S) (siehe Abb. 3A).

Die Sequenz wurde aufgeteilt in Fragemente, die den Hauptprozessierungsprodukten entsprachen: p83, p31, p36, p30 und p19. In den Übergangsregionen wurden Spalt-

stellen für Restriktionsendonukleasen (Pfeile in Abb. 3) so eingeplant, daß die Aminosäuresequenz nicht verändert wurde. Alle hier aufgeführten Schnittstellen kommen in der Sequenz nur einmal vor.

Die Fragmente wurden überlappend synthetisiert, so daß die Schnittstellen an den jeweiligen Enden die Verknüpfung zu den benachbarten Fragmenten durch Ligierung möglich machten. Alle Einzelfragmente enthielten zusätzlich an ihrem 5'-Ende eine BamHI-Schnittstelle zur Insertion in Expressionsvektoren. Über MluI und ClaI konnte die Gesamtsequenz kloniert werden. Das hier gezeigte Schema führt zunächst zu einer Sequenz, welche den GPI-Anker nicht ausbilden kann, da C-terminal 18 Aminosäuren fehlen. Die Synthese eines entsprechenden Oligonukleotids, sowie eines über die SphI-Schnittstelle sich erstreckenden "Primers" führt nach PCR zu dem Fragment GA, das über SphI und ClaI eingesetzt werden konnte, die resultierende Totalsequenz war gp190^S. Zur Entfernung der das Signalpeptid kodierenden Sequenz wurden "PCR-Primer" hergestellt, über die das Fragment ΔS synthetisiert wurde. Es erlaubt, über eine BamHI und eine HindIII-Schnittstelle den N-Terminus so zu verändern, daß das Protein mit Aminosäure Nr. 20 begann. Die Kernsequenz, welche das gp190/MSP1 ohne Signalsequenz und ohne GPI-Anker-Signal kodiert, wurde mit gp 190^{S2} bezeichnet. Deletion des GPI-Ankersignals allein führte zu gp190^{S1}.

B. Prinzip der PCR-gestützten Totalsynthese (siehe Abb. 3B)

Oligodesoxynukleotide von etwa 120 Nukleotiden Länge wurden abwechselnd vom kodierenden bzw. nichtkodierenden Strang so synthetisiert, daß sie jeweils etwa 20 Basen mit dem benachbarten Fragment überlappten. Das Schema zeigt beispielhaft die Synthese eines ca. 800 bp langen Fragments aus Oligonukleotiden. In der ersten Stufe wurden in 4 Reaktionsgefäßen jeweils 2 Oligonukleotide "asymmetrisch" amplifiziert. Es entstanden 4 etwa 220 bp lange DNA-Populationen, die vorwiegend aus Einzelsträngen bestanden (A, B, C, D). Vereinigung von A und B sowie von C und D und Amplifikation über 5 Zyklen führte zu 2 etwa 400 bp langen doppelsträngigen Produkten. Asymmetrische Amplifikation dieser DNA-Fragmente (Stufe III) ergab Einzelstrangpopulationen, welche nach Vereinigung und Amplifikation (Stufe IV) nach

10 Zyklen das Endprodukt G von ca. 800 bp Länge ergaben. Diese Synthese war ohne Isolierung der Zwischenprodukte und ohne Puffer oder Enzymerneuerung durchführbar, und war nach 3 Stunden beendet. Das Endprodukt wurde elektrophoretisch gereinigt, mit den geeigneten Restriktionsendonukleasen nachgeschnitten und im pBluescript (Stratagene), in dessen Polylinker eine MluI und eine ClaI-Schnittstelle eingesetzt worden waren, in *E. coli* kloniert.

C. Totalsequenz des gp190^S (siehe Abb. 3C)

Nach Fusion aller Teilsequenzen (Abb. 3A) in pBluescript wurde die Sequenz des Gens mit der Dideoxymethode verifiziert. Das Leseraster des gp190^S hatte eine Länge von 4917 bp (+2 Stopcodons) und kodierte eine Aminosäuresequenz, welche der des gp190/MSP1 aus FCB-1 entspricht (1639 Aminosäuren).

D. N- und C-Terminus der gp190^{S1}-Variante (siehe Abb. 3D)

Die N-terminale Sequenz, beginnend mit der BamHI-Schnittstelle, zeigte den Übergang bei Aminosäure 20, von der angenommen wird, daß sie nach Abspaltung des Signalpeptids den N-Terminus definiert. Am C-Terminus war die kodierte Sequenz bei Aminosäure 1621 zu Ende. Den Stopcodons folgte die ClaI-Schnittstelle.

Beispiel 2: Expression des gp190^{S2} in *E. coli*

A. Expressionsvektor (siehe Abb. 4A)

Die gp190^{S2}-Sequenz wurde über die BamHI- und ClaI-Schnittstellen in pDS56RBSII eingesetzt. Dadurch wurden 6 Histidine sowie einige aus dem Vektor stammenden Aminosäuren an den N-Terminus fusioniert; dies ergibt folgende N-terminale Sequenz des Leserasters; Met Arg Gly Ser (His)₆ Gly Ser. Durch den Promotor P_{N25lacO-1} stand die Transkription unter lacR/O/IPTG-Kontrolle.

B. Expression und Aufreinigung von gp190^{S2} (siehe Abb. 4D)

Eine Überführung des Vektors pDS56RBSIIgp190^{S2} in *E. coli* DH5αZ1 und Induktion der Synthese durch IPTG ergab nach elektrophoretischer Auftrennung des Protein-Totalextraktes der Kultur eine deutlich sichtbare Bande in der erwarteten Größe

(Pfeil). Die Aufreinigung des Materials durch IMAC und Affinitätschromatographie (Antikörpersäule mit mAK5.2) führte zu einem homogenen Produkt von etwa 190 kD. In der Abb. bedeuten M= Molekulargewichtsstandards; 1= *E. coli* Proteine vor, 2= nach Induktion mit IPTG für 2 Stunden. 3,4,5 = Fraktionen aus der Elution der mAK-Säule.

Beispiel 3: Tetrazyklin-kontrollierte Expression von gp190^{S1} in HeLa- und CHO-Zellen und Isolation des Produktes (siehe auch Abb. 5) bzw. 6c)

A. Die gp190-Sequenz wurde in den Expressionsvektor pBi-5 über die BamHI/Clal-Schnittstellen eingesetzt. Damit stand die Transkription des Gens unter Kontrolle eines bidirektionalen "tTA-responsive"-Promotors und konnte über Tc reguliert werden. Der bidirektionale Promotor initiierte gleichzeitig die Transkription des Indikatorgens Luciferase. Damit ließ sich die Regulation der Expression leicht verfolgen (siehe auch Abb. 5A).

B. Immunfluoreszenz von HeLa-Zellen, welche Luciferase und gp190^{S1} Tc-kontrolliert exprimieren.

In HtTA93-9-Zellen, welche die bidirektionale Transkriptionseinheit von (A) enthalten, wurde mit Antikörpern die Produktion von Luciferase (links), gp190^{S1} (Mitte) in Abwesenheit von Tc nachgewiesen. Nach Zugabe von Tc zeigte sich keine nennenswerte Synthese von gp190^{S1}, (wie dargestellt in Abb 5B, rechts).

C. Elektrophoretische Charakterisierung von aus HeLa-Zellen aufgereinigtem gp190^{S1}.

HeLa-Zellklon HtTA93-9 sowie CHO-Zellklon CHO27-29 wurden mit oder ohne Tc kultiviert. Durch Elektrophorese aufgetrennte Zellextrakte wurden mittels "Western blot" mit mAK5.2 analysiert (Abb. 5C); links zeigt eine Analyse der CHO-, rechts der HeLa-Zelllinie. (1) = Kultur ohne, (2) = Kultur mit Tc, (3) = nicht transfizierte HtTA-1-Zelllinie, Molekulargewichtsstandards sind jeweils links angedeutet.

D. Aufreinigung von in HeLa-Zellklon HtTA93-9 synthetisiertem gp190^{S1}

Die präparative Aufzucht der HtTA-93-9-Linie und Induktion der Expression von der gp190^{S1} durch Tc-Entzug erlaubte die Isolierung des Genproduktes über Affinitätschromatographie (mAK5.2-Säule).

Das Coomassie gefärbte Polyacrylamid-Gel (Abb. 6C) zeigte nach Elektrophorese ein Produkt, das aus gp190^{S1} sowie einem weiteren Protein von ca 50 kD bestand. Letzteres war kein Derivat von gp190^{S1}, stammte also aus HeLA-Zellen. Seine gezielte Abtrennung sollte jedoch keine prinzipielle Schwierigkeit darstellen.

Beispiel 4: Expression von gp190^S in Toxoplasma gondii und Aufreinigung des Produktes (siehe hierzu auch Abb. 6).

A. Die gp190^S-Sequenz wurde über MluI/PstI in den Vektor ppT eingesetzt. Damit wurde das Gen unter die Kontrolle des Tubulin-Promotors (P_{tub-1}) von T. gondii gebracht. Die 3' nicht-translatierte Region (VTR) stammte von dem Hauptoberflächenprotein von T. gondii (SAG-1) ab.

B. Expression von gp190^S in T. gondii.

Transfektion von T. gondii mit pTT190 führte zur Isolierung von Parasitenlinien, die konstitutiv gp190^S exprimierten. Die Immunfluoreszenz mit mAK5.2 (mittleres Bild) zeigte nicht nur die Expression des Gens, sondern legte auch die Verankerung des Expressionsproduktes an der Oberfläche des Parasiten nahe, da es, wie SAG-1, das gleiche Immunfluoreszenz-Muster erzeugte (rechter Teil der Abb. 6B); links in Abb. 6B ist eine Phasenkontrastaufnahme des mittleren Bildes dargestellt.

C. Isolierung von gp190^S aus T. gondii.

Aus präparativen Mengen von T. gondii (5x10⁹ Parasiten) wurde gp190^S mittels Affinitätschromatographie (mAK5.2-Säule) aufgereinigt. Das hochreine Protein besaß das erwartete Molekulargewicht, wie das Coomassie-gefärbte Polyacrylamid-Gel nach Elektrophorese zeigte (2-3 aus Abb. 6C). Bei Nr. (1) Abb. 6C ist gereinigtes gp190^{S1} aus CHO-Zellen dargestellt mit Molekulargewichtsmarkierung an der linken Seite.

Beispiel 5: Charakterisierung des gp190^S mit monoklonalen Antikörpern.

Die Wechselwirkung von 16 monoklonalen Antikörpern mit gp190^S aus den verschiedenen heterologen Expressionssystemen wurde durch Immunfluoreszenz (IFA) an *P. falciparum* und *T. gondii* bzw. durch "Western blot" an den aufgereinigten Proteinen überprüft. Völlige Übereinstimmung wurde gefunden, wenn die beiden Parasiten verglichen wurden (Zahl der + deutet die relative Intensität der Fluoreszenz an). Im Western blot reagieren 12 mAK's mit gp190^S aus *E. coli* und *T. gondii*. Im Gegensatz dazu binden 3 Antikörper nicht an das aus CHO-Zellen isolierte Material. Antikörper 15 und 16, die Epitope aus dem oligomorphen bzw. dem alternativen Allel (MAD20) erkennen, reagieren nicht. Die Ergebnisse sind in Tab. 1 zusammengefaßt, wobei ND = nicht durchgeführt bedeutet.

Beispiel 6: Expression des gp190^S in heterologen Systemen.

1. Expression in *E. coli*

Das gp190^{S2} wurde in den Expressionsvektor, pDS56, RBSII eingesetzt, wo es unter Kontrolle des P_{N25lacO-1}-Promotors stand, der über das lac Operator/Repressor/IPTG-System regulierbar ist (Abb. 4A). Die Überführung des Plasmids in Repressor-produzierende *E. coli*-Zellen, z.B. *E. coli* DH5 α Z1 erlaubt es, das gp190^{S2} unter IPTG-Kontrolle zu exprimieren. Das Produkt war aus dem Rohextrakt mittels einer Nickel-Chelatsäule über die durch den Vektor eingebrachte N-terminale (His)₆-Sequenz isolierbar. Eine anschließende Affinitätschromatographie an einer Antikörpersäule führte zu einem hochreinen Präparat. Da der verwendete monoklonale Antikörper (mAk5.2) ein konformationelles Epitop im C-terminalen Bereich erkannte, wird durch diese 2-Stufenreinigung auf intaktes Protein voller Länge, mit korrekter Faltung zumindest am C-Terminus, selektioniert (Abb. 4B).

Das Endprodukt besitzt, im Gegensatz zu dem natürlichen Material, am N-Terminus 11 zusätzliche Aminosäuren, davon 6 Histidine. Es enthält keine N-terminale Signal- und auch keine C-terminale Anker-Sequenz. Die *P. falciparum*-spezifische Sequenz beginnt mit Aminosäure 20 und endet mit Aminosäure 1621.

2. Kontrollierte Expression des gp190^{S1} in HeLa- und CHO-Zellkulturen.

Das gp190^{S1} wurde in den Vektor pBi-5 eingesetzt und damit unter Kontrolle eines durch Tetrazyklin (Tc) regulierbaren Promotors gestellt. Das Tc-kontrollierte System wurde aus 2 Gründen gewählt:

- Es gehört zu den Expressionssystemen, mit denen höchste Ausbeuten in Säugerzellen erreicht werden.
- Nicht-sekretierte Fremdproteine in hoher Konzentration können mit dem Metabolismus der Zellen in negativer Weise interferieren. Die Synthese des gewünschten Produktes wird daher erst nach Aufwachsen der Kultur initiiert.

In dem Konstrukt pBi5-gp190^{S1} wurde ein bidirektionaler Promotor durch Tc-kontrollierten Transkriptionsaktivator (tTA) aktiviert und initiierte die Transkription sowohl des gp190^{S1} als auch des Luziferase-Indikatorgens. In Anwesenheit von Tc ist der Promotor inaktiv. Die Transkriptionseinheit wurde sowohl in HeLa als auch in CHO-Zellen, welche beide konstitutiv tTA synthetisieren (HtTA-1-Linie (Gossen, M. and Bujard, H. (1992), Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA 89, 5547-5551); CHO-tTA-Linie, unveröffentlicht), überführt. Durch Kōtransfektion (Ca²⁺-Phosphat-Methode) mit einem Hygromycin-Resistenz-vermittelnden Markergen wurde auf erfolgreiche chromosomale Integration selektioniert. Hygromycin-resistente Klone wurden dann auf Regulierbarkeit der Expression \geq Tc untersucht, indem die Luziferaseaktivität als Indikator genutzt wurde. In gut regulierbaren Klonen (Regulationsfaktor \leq Tc 1000) wurde die gp190 Synthese geprüft. Immunfluoreszenz-Analyse (Abb. 5B) sowie Untersuchungen über "Western blot" (Abb. 5C) erlaubten, von beiden Zelltypen Klone zu identifizieren, welche gp190 unter streng regulierbaren Bedingungen synthetisieren. Von 20 Klonen wurde jeweils der bestregulierbare subkloniert. Die Subklone HtTA93-9 sowie CH027-29 wurden für Kulturen im 10 l Maßstab verwendet. Aus Zellextrakten dieser Kulturen ließ sich intaktes gp190^{S1} mittels Affinitätschromatographie (mAk5.2) isolieren. Das Material ist homogen bis auf eine einzige zelluläre Komponente, die nicht von gp190^{S1} abstammt und die etwa 25% des Präparats ausmacht (Abb. 6C). Sie müßte in einem weiteren Reinigungsschritt entfernt werden.

3. Expression des gp190^S in Toxoplasma gondii.

Toxoplasma gondii gehört wie *P. falciparum* zu den Apicomplexa und hat daher wahrscheinlich ein dem *P. falciparum* ähnliches Protein-Modifikationssystem. *T. gondii* läßt sich mit Fremd-DNA transfizieren, die effizient in das Genom integriert wird, außerdem läßt sich *T. gondii* problemlos in Zellkultur vermehren. Um ein möglichst natives gp190-Produkt zu erhalten, wurde gp190^{S2} so exprimiert, daß das Protein sekretiert und auf der Oberfläche des Parasiten, wie bei *P. falciparum*, über ein GPI-Motiv in der Membran verankert wird. Dazu wurde das gp190^{S2} (Abb. 3A) in das Plasmid ppTMCS (D. Soldati, unveröffentlicht) eingesetzt (Abb. 6A) und damit unter Kontrolle des *T. gondii*-Tubulin Promotors gestellt.

Dieses Expressionskonstrukt wurde in *T. gondii* transfiziert. Selektion mit Chloramphenicol führte zu resistenten Klonen, die gp190 synthetisieren, wie durch Immunfluoreszenz nachgewiesen wurde (Abb. 6B). Die Immunfluoreszenz mit anti-gp190-Antikörpern war nicht unterscheidbar von einer entsprechenden Anfärbung der Parasiten mittels Antikörper gegen SAG1, dem Hauptoberflächenprotein von *T. gondii*. Es ist daher davon auszugehen, daß gp190 an der Oberfläche von *T. gondii* verankert ist. Mehrere *T. gondii*-Klone (Nr. 3.1 bis 3.4) wurden charakterisiert und für die Produktion von gp190 aufbewahrt. Aus in präparativem Maßstab aufgewachsenen *T. gondii*-Kulturen (Klon 3.4) wurde gp190 mittels Affinitätschromatographie (mAK5.2.-Säule) isoliert. Analyse im elektrischen Feld zeigte ein homogenes Produkt mit einer Wanderungsgeschwindigkeit, die auf das intakte Protein schließen läßt (Abb. 6C).

Beispiel 7: Charakterisierung von gp190 Protein aus verschiedenen Expressionssystemen mittels monoklonaler Antikörper.

Ein Satz gp190-spezifischer monoklonaler Antikörper, von denen mehrere konformationelle Epitope erkennen, wurde dazu benutzt, über Immunfluoreszenz die Reaktivität der Antikörper mit *P. falciparum*- bzw. *T. gondii*-Parasiten zu vergleichen. Tabelle 1 zeigt, daß die Reaktivität der 16 Antikörper für beide Parasiten gleich ist. Dies ist ein starker Hinweis darauf, daß in *T. gondii* weitestgehend "natives" gp190 produziert wird. Der Vergleich der Reaktivität der Antikörper mit Protein aus *E. coli*, HeLa- bzw. CHO-Zellen sowie *T. gondii* zeigt ebenfalls, daß die meisten Antikörper mit den 4 Präparaten reagieren. Insbesondere wird das aus *E. coli* isolierte Protein von mehr

Antikörpern erkannt als das Produkt aus Säugerzellen. Dies ist wahrscheinlich eine Konsequenz der Glykosylierung in Säugerzellen.

Beispiel 8: Immunsierung von Aotus lemurinus griseimembra-Affen mit gp190/MSP1 aus *P. falciparum* (FCB-1).

Zwei unabhängige Immunsierungsexperimente (A, B) wurden durchgeführt. Dazu wurde aus jeweils ca. 2×10^{11} Parasiten unter schonenden Bedingungen einmal 1,0 mg (A) und einmal 0,6 mg hochreines gp190/MSP1 isoliert.

Das Protein wurde zusammen mit Freund'schem Adjuvans (FCA) verabreicht. Die Kontrollgruppe erhielt lediglich FCA. Es wurde dreimal im Abstand von 4 Wochen mit gleicher Menge an Protein bzw. mit Adjuvans immunisiert. Zwei Wochen nach der letzten Immunisierung wurden die Tiere mit jeweils 10^5 Parasiten (FVO-Stamm) aus einem Donor-Tier infiziert. Die Parasitämie wurde täglich gemessen. Die Ergebnisse sind in Abb. 2 zusammengefaßt. Dabei bedeutet

T: daß die Tiere mit Resochin behandelt wurden

D: ein verstorbene Tier

Abb. 2A.: die Individuen der geimpften Gruppe erhielten je 3 x 60 Mikrogramm gp190/MSP1

Abb. 2B: die Individuen der geimpften Gruppe erhielten je 3 x 40 Mikrogramm gp190/MSP1

Während in den Kontrollgruppen nur 1/11 Tiere keine Parasitämie entwickelten, waren es in der geimpften Gruppe 6/10. Die vier Tiere aus der geimpften Gruppe, die eine hohe Parasitämie entwickelten, taten dies - im Vergleich zur Kontrollgruppe - mit einer durchschnittlichen Verzögerung von vier Tagen (Überschreiten der 2%-Grenze der Parasitämie).

Diese Experimente zeigten erstmals einen hochsignifikanten Schutz gegen Infektion mit *P. falciparum* durch gp190/MSP1 im Affenmodell. Das erfindungsgemäße Verfahren erlaubt es somit erstmalig, einen wirksamen Impfstoff gegen die Malaria anzugeben.

Tabelle 1: Wechselwirkung von gp190^S mit monoklonalen Antikörpern

Code	mAb	Art des Epitops	Variabilität	IFA			Western blot	
				P.f. FCB	Toxoplasma	E. coli	Toxo-plasma	CHO
1	5,2	konformationell	konserviert	++++	++++	+	+	+
2	12,10	konformationell	konserviert	++++	++++	+	+	+
3	7,5	konformationell	konserviert	++++	++++	+	+	+
4	12,8	konformationell	konserviert	++	++	+	+	+
5	7,3	konformationell	dimorph (K1)	++++	+++	+	+	+
6	2,2	konformationell	konserviert	++++	++++	+	+	+
7	7,6	konformationell	dimorph (K1)	++++	++++	+	+	+
8	9,8	konformationell	konserviert	++++	++	+	+	-
9	13,2	sequentiell	konserviert	++++	++++	+	+	+
10	13,1	sequentiell	dimorph (K1)	++++	+++	+	+	-
11	6,1	sequentiell	dimorph (K1)	++++	++++	+	+	ND
12	A5Z	nicht bekannt	nicht bekannt	+++	+++	+	+	+
13	17,2	nicht bekannt	nicht bekannt	++++	+++	ND	ND	ND
14	15,2	nicht bekannt	nicht bekannt	++++	+++	ND	ND	ND
15	9,7	konformationell	dimorph (MAD20)	-	-	-	-	-
16	12,1	sequentiell	oligomorph	-	-	-	-	-

Patentansprüche

1. Herstellungsverfahren für das vollständige gp190/MSP1-Protein von Plasmodium, insbesondere *Plasmodium falciparum*, **dadurch gekennzeichnet**, daß das vollständige Gen für gp190/MSP1 in einem geeigneten System, vorzugsweise einem Wirtorganismus, exprimiert wird.
2. Rekombinantes Herstellungsverfahren gemäß Anspruch 1, **dadurch gekennzeichnet**, daß der Synthese der dem Protein zugrunde liegenden DNA-Sequenz die DNA-Sequenz des *P. falciparum*-Stammes FCB-1 zugrunde gelegt wird.
3. Rekombinantes Herstellungsverfahren gemäß Anspruch 2, **dadurch gekennzeichnet**, daß der AT-Gehalt der exprimierten DNA-Sequenz gegenüber der natürlich vorkommenden Sequenz reduziert wurde, vorzugsweise von 74% auf 55%.
4. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz einschließlich Signalpeptid und Ankersignal kodiert.
5. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals kodiert.
6. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals und des Signalpeptides kodiert.
7. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 6, **dadurch gekennzeichnet**, daß es folgende Schritte umfaßt:

- (a) Entwurf der zu synthetisierenden DNA-Sequenz aus *P. falciparum* FCB-1, wobei eine DNA-Sequenz mit den im menschlichen Genom üblichen Codon-Häufigkeiten unter Erhalt der Aminosäuresequenz des FCB-1 Proteins hergestellt werden sollte,
 - (b) Einteilung der entworfenen Sequenz in überlappende Regionen, vorzugsweise in Regionen p83, p31, p36, gp30 und gp19,
 - (c) Synthese von Desoxyoligonukleotiden, die jeweils mindestens die gesamte Länge einer Region abdecken,
 - (d) Synthese der kodierenden Regionen für gp19, gp30, p36 und p31 durch PCR und Synthese der kodierenden Region für p83 durch Fusion aus zwei etwa 1200bp umfassenden Sequenzen,
 - (e) einzelne Klonierung der kodierenden Sequenzen
 - (f) Fusion des gesamten Gens und
 - (g) Expression in einem geeigneten Expressionssystem.
8. Rekombinantes Herstellungsverfahren nach Anspruch 7, **dadurch gekennzeichnet**, daß die in Schritt (c) synthetisierten Desoxyoligonukleotide durchschnittlich 120 Nukleotide lang sind und die benachbarten Sequenzen jeweils um ca 20 Basen überlappen.
9. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor dPS56, RBSII verwendet wird.
10. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor pBi-5 verwendet wird.
11. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, daß als Expressionsvektor ppTMCS verwendet wird.

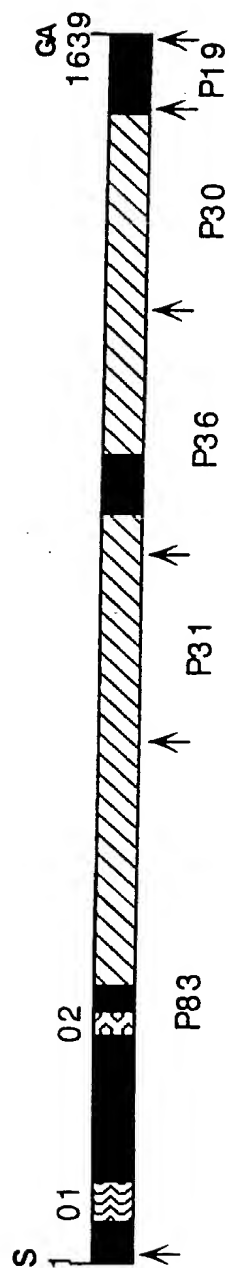
12. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in *E. coli* exprimiert wird.
13. Rekombinantes Herstellungsverfahren nach Anspruch 12, **dadurch gekennzeichnet**, daß der verwendete *E. coli*-Stamm der Repressor-produzierende Stamm *E. coli* DH5alphaZ1 ist.
14. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in HeLa-Zellen exprimiert wird.
15. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in CHO-Zellen exprimiert wird.
16. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, **dadurch gekennzeichnet**, daß die dem Protein zugrunde liegende DNA-Sequenz in *Toxoplasma gondii* oder *Leishmania* exprimiert wird.
17. Vollständige, zur Expression geeignete DNA-Sequenz des gp190/MSP1-Oberflächenproteins von Plasmodium, insbesondere *P. falciparum*, vorzugsweise erhältlich durch das rekombinante Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 16.
18. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 17, **dadurch gekennzeichnet**, daß sie nicht für das Ankersignal kodiert.
19. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 18, **dadurch gekennzeichnet**, daß sie weder für das Ankersignal noch für das Signalpeptid kodiert.

20. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 19, **dadurch gekennzeichnet**, daß sie eine Sequenz für am N-Terminus vorliegende 11 zusätzliche Aminosäuren, davon 6 Histidine, umfaßt.
21. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 20, **dadurch gekennzeichnet**, daß die Sequenz keine erkennbaren "splice donor"- und "splice acceptor"-Signale enthält.
22. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 21, **dadurch gekennzeichnet**, daß die Sequenz keine größeren GC-reichen Sequenzen enthält.
23. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 22, **dadurch gekennzeichnet**, daß die Sequenz keine Erkennungssignale für Restriktionsenzyme, welche Sequenzen von sechs oder mehr Basenpaaren erkennen, enthält.
24. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 23, **dadurch gekennzeichnet**, daß die Sequenz für Erkennungssignale bestimmter Restriktionsnukleasen in Regionen, die die nach der Prozessierung des Proteins entstehenden Domänen trennen, einmal vorkommende Schnittstellen für Restriktionsendonukleasen enthält.
25. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 24, **dadurch gekennzeichnet**, daß die Sequenz an ihren beiden Enden Schnittstellen für Restriktionsendonukleasen aufweist, die in der übrigen Sequenz und in einem zu verwendenden Vektor nicht vorkommen.
26. Wirtsorganismus, der die vollständige Nukleinsäuresequenz für das gp190/MSP1-Oberflächenprotein und/oder das vollständige Protein enthält.
27. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus *E. coli* ist.

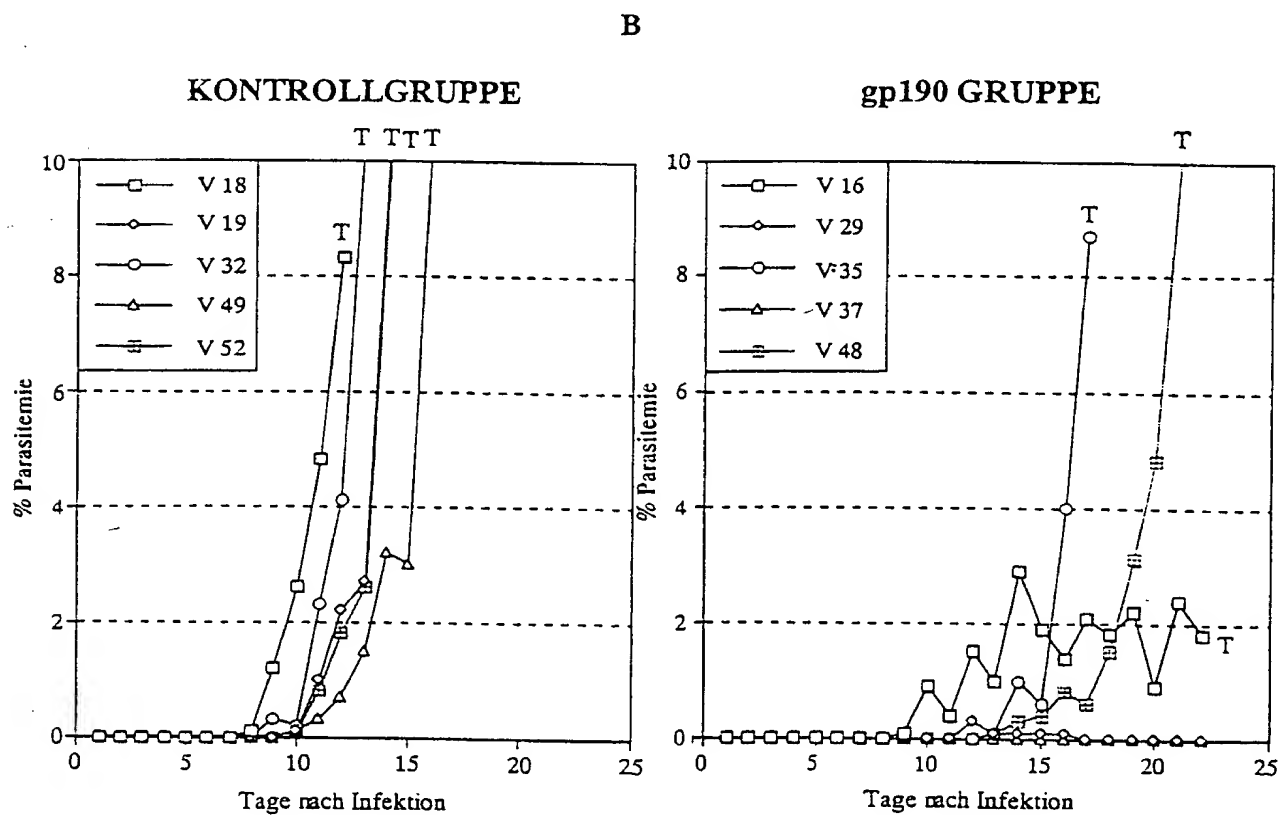
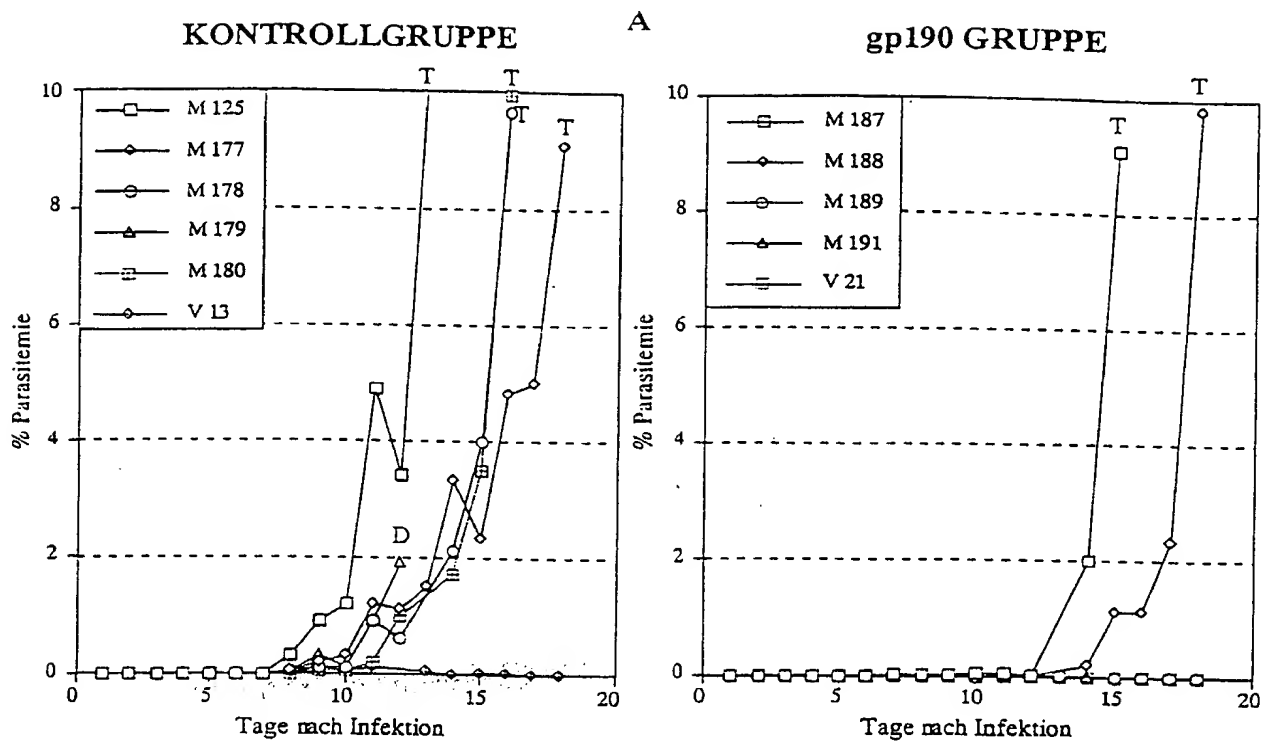
28. Wirtsorganismus gemäß Anspruch 27, **dadurch gekennzeichnet**, daß der *E. coli*-Stamm der Repressor-produzierende *E. coli*-Stamm DH5alphaZ1 ist.
29. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus HeLa-Zellen sind.
30. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus CHO-Zellen sind.
31. Wirtsorganismus gemäß Anspruch 29 oder 30, **dadurch gekennzeichnet**, daß die Wirtszellen konstitutiv tTA synthetisieren.
32. Wirtsorganismus gemäß Anspruch 26, **dadurch gekennzeichnet**, daß der Wirtsorganismus *Toxoplasma gondii*, *Leishmania*, Baculoviren, Adenoviren oder Hefen vorgesehen sind.
33. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur aktiven Immunisierung gegen Malaria.
34. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur Herstellung von monoklonalen, zur passiven Immunisierung geeigneten Antikörpern.
35. Verwendung einer gemäß der Ansprüche 1 bis 16 hergestellten DNA-Sequenz zur Herstellung einer Vakzine auf Nukleinsäurebasis.
36. Verfahren zur Stabilisierung von Gen-Sequenzen, **dadurch gekennzeichnet**, daß der AT-Gehalt der Sequenz verringert wird.
36. Stabilisiertes Gen, **dadurch gekennzeichnet**, daß es einen geringeren AT-Gehalt aufweist als das nicht stabilisierte Gen.

37. Vektor enthaltend eine DNA-Sequenz nach einem der Ansprüche 17 bis 25 und/oder 36.
 38. Wirtszelle enthaltend einen Vektor nach Anspruch 37.
 39. Impfstoff enthaltend ein Protein, hergestellt nach einem Verfahren nach einem der Ansprüche 1-16 und/oder eine DNA-Sequenz nach einem der Ansprüche 17-25 und/oder einen Wirt nach einem der Ansprüche 26-32 und/oder einen Vektor nach Anspruch 37.
 40. Impfstoff nach Anspruch 39, dadurch gekennzeichnet, daß er weitere Immunität hervorrufende Produkte aus Plasmodium, insbesondere *P. falciparum*, enthält.
-

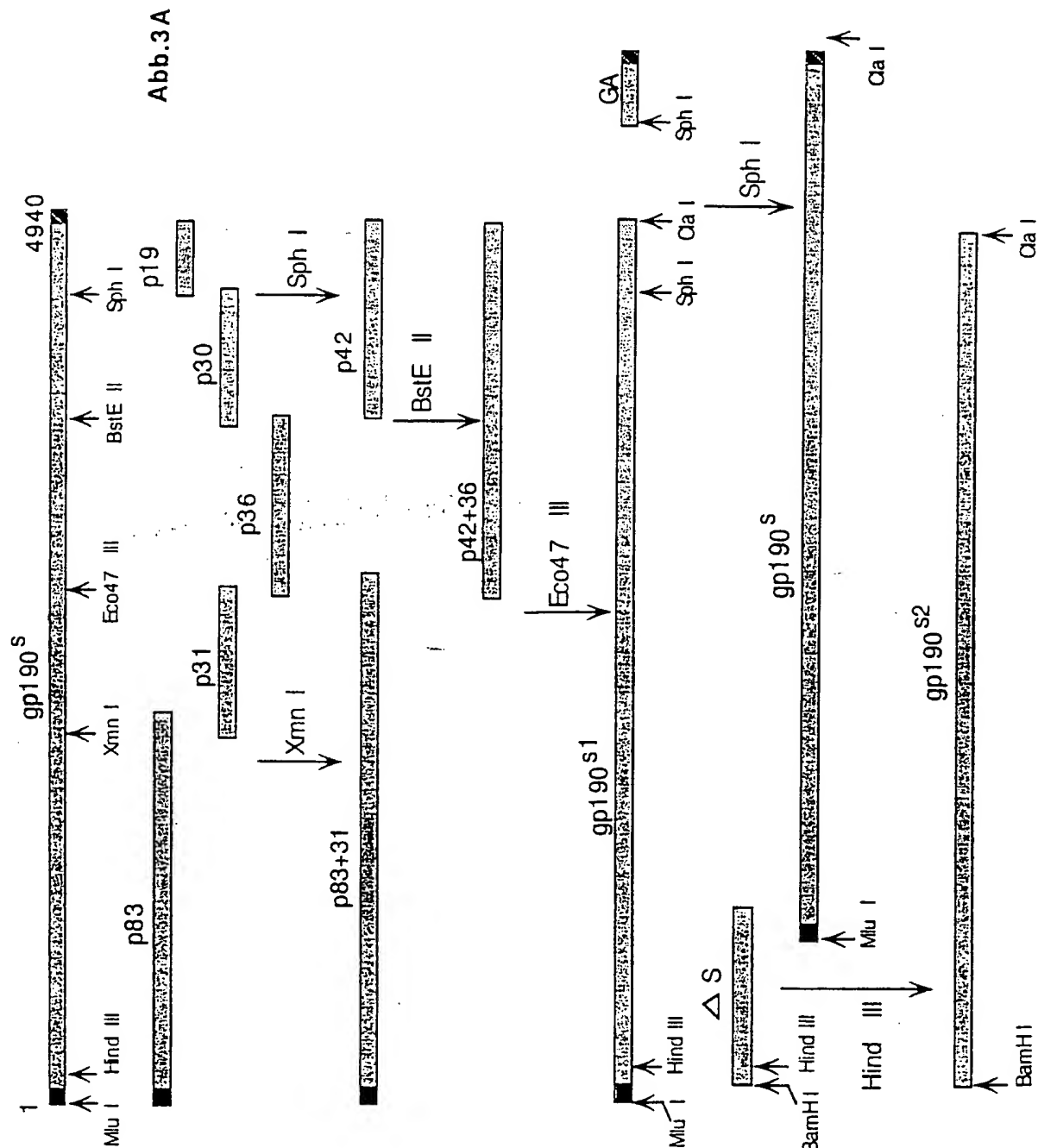
Abb.1



THIS PAGE BLANK (USPTO)

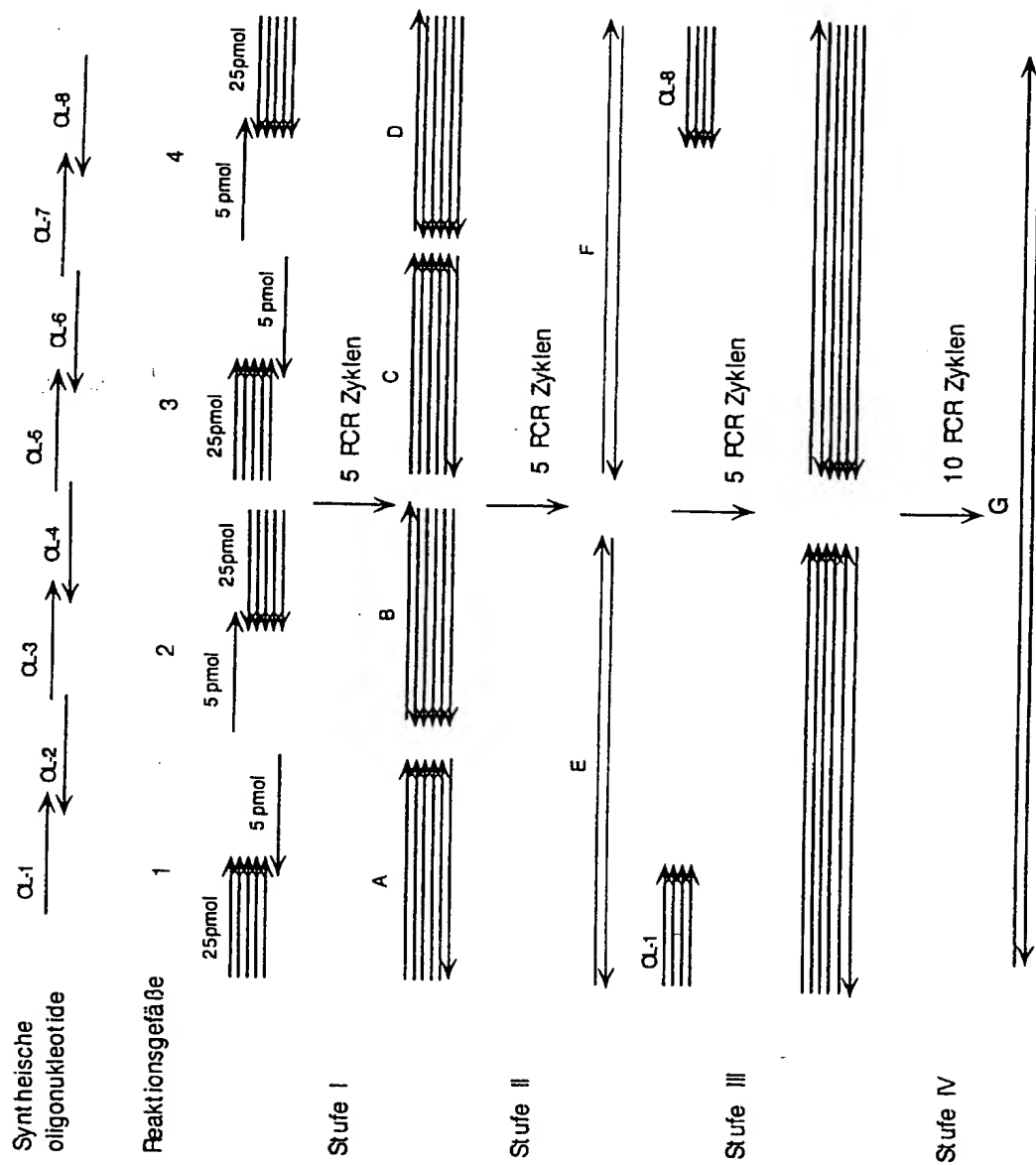


THIS PAGE BLANK (USPTO)



THIS PAGE BLANK (USPTO)

Abb.3B



THIS PAGE BLANK (USPTO)

Abb 3c

5/16

**DNA Sequenz des nativen(gp190ⁿ) und des
synthetischen(gp190^s) Gens für gp190 aus FCB-1**

```

AS      M K I I F F L C S F L F F I I N T Q C V T H E S Y Q E      27
gp190n      G A T T A T T A A A T A A T A A T A A T A A      30
gp190s CGCAGCGGTATGAAAAATCATTTTCTCTCTGTTCTTTTATCATCAATACTCAGTCCGTGACCCAGCAATCCTATCAGGAG      90
      M I I I

AS      L V K K L E A L E D A V L T G Y S L F Q K E K M V L N E G T      57
gp190n      T C A A A A A T G A T T T T A T A A A A T A A A A      60
gp190s CTGGTTAAGAAACTGGAAGCTTTGGAAGATGCCGTCTTACCGGATACAGCCTGTTCCAGAAGGAGAAGATGGTCTGAATGAAGGGACG      180

AS      S G T A V T T S T P G S K G S V A S G G S G G S V A S G G S      87
gp190n      A A T T T T A G T A T T C A T A C A T T A T C A      90
gp190s AGTGGCAGCGCGTTACACCAGCACACCCGGTCTAAAGGGTCTGTGGCTAGCGGTGGCTCCGGTGGGTCTGTGGCTCTGGGGGTTC      270

AS      V A S G G S V A S G G S V A S G G S G N S R R T N P S D N S      117
gp190n      T T A T T C A T T T T T C A T A T T C A A C T A T A T T A      120
gp190s GTCCCTCCGGCGGACGCGTGGCATCAGTGGCTCAGTGGCAAGCGCGGTTCGGGAACAGTCAAGAACCAATCCATCTGACAACTCT      360

AS      S D S D A K S Y A D L K H R V R N Y L L T I K E L K Y P Q L      147
gp190n      T A T T A T T T A A A A C T C T G T A A A C A T T A C C      150
gp190s AGCGATTCCGACGCCAAGTCTACGCGGACCTCAAGCAGCGAGTGAGAAATATCTCTCACTATCAAGGAGCTGAAGTACCCACAGTTG      450

AS      F D L T N H M L T L C D N I H G F K Y L I D G Y E E I N E L      177
gp190n      T T T A T A T T T T T T T T T T T T T T T T T T A T A      180
gp190s TTCGACCTCACTAATCATATGCTGACACTGTGTGATAACATTCATGGCTTCAAATATCTGATTGACGGTTACGAAGAGATCAATGAATC      540

AS      L Y K L N F Y F D L L R A K L N D V C A N D Y C Q I P F N L      207
gp190n      T A T A A C T T T T T A A A T A T A T T T T T T T T T C T      210
gp190s CTGTACAAAGTTGAATTTCTACTTCGACTTGCTAAGGGCCAACTGAATGACGTTTGGCGCAATGACTATTGTCAAATTCATCAATTTG      630

AS      K I R A N E L D V L K K L V F G Y R K P L D N I K D N V G K      237
gp190n      A T C T A T A A C T A A C T G A A A A T A T A T T A T A A      240
gp190s AAGATCAGAGCCACGAGTTGGAGCTATTGAAGAAGTTGGTCTTCGGATATCGCAAGCCTCTCGACAACATCAAGGCAATGTGGGAAAG      720

AS      M E D Y I K K N K K T I E N I N E L I E E S K K T I D K N K      267
gp190n      C A A A A A T A T A T A T A T A T A T A T A T A T T      270
gp190s ATGGAAGATTATATATAAAGAATAAGAAGACCATCGAGAACATTAAACGAGCTGATCGAAGAATCCAAAAGACCATAGACAAAAATAAG      810

AS      N A T K E E E K K K L Y Q A Q Y D L S I Y N K Q L E E A H N      297
gp190n      T A A A A A A A T A T T T T T T T T T T T T T A T A T      300
gp190s AATGCAACCAAGGAGGAAGAAAGAAGAGTTGTACCAGGCCAGTACGACCTGTCCATCTATAACAACAGCTTGAAGAAGCCCATAC      900

AS      L I S V L E K R I D T L K K N E N I K E L L D K I N E I K N      327
gp190n      T A A T T A A A T T T T A A A C T G T A T T T A A A A      330
gp190s CTCATCAGCGTACTGGAGAAGCGCATAGACACCTCAAGAAGAAATGAAATATCAAAGAAGCTGCTCGACAAGATTAAATGAAATTAAGAAT      990

AS      P P P A N S G N T P N T L L D K N K K I E E H E K E I K E I      357
gp190n      C A G A T A T A A T T C T T A A C A A C A A A A A A A T      360
gp190s CCTCCGCGACCAACTCTGGGAACACCCCTAACACGCTCTGGACAAGAACAAGATAGAGGAGCAGCAAGAGAGATCAAGAGATC      1080

AS      A K T I K F N I D S L F T D P L E L E Y Y L R E K N K N I D      387
gp190n      T A T T T A G T A A A T A A T A A A A A A A T T      390
gp190s GCCAAAACCATTAAGTTCAACATAGATTCTCTTTACTGATCCCTTGAGCTGGAGTACTACTTGAGAGAGAAGAATAAGAATATAGAC      1170

AS      I S A K V E T K E S T E P N E Y P N G V T Y P L S Y N D I N      417
gp190n      A A G T A G T A T C A T C A A T G A A T A T C C A A T G G T G T G A C T A C C C T C T G T C T T A T A A C G A T A T C A A C      420
gp190s ATCTCCGCCAAAAGTCGAGACAAAGGAATCAACCGAACCTTAATGAATATCCCAATGGTGTGAGCTACCCCTCTGTCTTATAACGATATCAAC      1260

AS      N A L N E L N S F G D L I N P F D Y T K E P S K N I Y T D N      447
gp190n      T T A T A T T C T T T A T A T A T A A A A G A C A T T T      450
gp190s AACGCTCTCAACGAGCTCAATAGCTTCGGTGACTTGATTAAACCCCTTCGATTATACGAAAGAACCCTCTAAGAATATCTACACAGACAAT      1350

AS      E R K K F I N E I K E K I K I E K K K I E S D K K S Y E D R      477
gp190n      A A A C A T T A A T A A A A A T C T A T C A A A A A A A      480
gp190s GAGAGAAAGAGTTTATCAACGAAATCAAGGAGAAGATCAAAATTCAGAGAAGAAATTTGAGAGTGACAAAGAAAGTTACGAAGACCGC      1440

AS      S K S L N D I T K E Y E K L L N E I Y D S K F N N N I D L T      507
gp190n      T C T G T C T T A A A A A T A T A T A T A G T A T T A T T A T      510
gp190s AGCAAAAGTCTAAACGATATCACTAAAGAGTATGAAAGCTGCTGAACGAGATCTATGATTCCAAATTCACAAATAACATCGACCTGACC      1530

AS      N F E K M H G K R Y S Y K V E K L T H H N T F A S Y E N S K      537
gp190n      T A T A A T A T T T T T T T T T T T T T T T T A A A      540
gp190s AACTTCGAGAAAATGATGGGAAAACGGTACTCTTACAAAGTGGGAAAAGTACACACCATAATACCTTTGTCATCTATGAGAATTCTAAG      1620

```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

AS	L E K K K L S Y L S S G L H H L I A E L K E V I K N K N Y T	1167
gp190 ⁿ	T A A A A T A T C A T A A T T A T T A T T A T A A A A A T A T T A	
gp190 ^s	CTGGAGAAGAAGAGCTCAGCTACCTCTCTAGCGGACTGCATCACCTGATCGCCGAGCTCAAGGAAGTCATTAGAACAAGAATACACC	3510
AS	G N S P S E N N T D V N N A L E S Y K K F L P E G T D V A T	1197
gp190 ⁿ	T T C T T A G T T C T T A A A T C A T A A	
gp190 ^s	GGCAATAGCCCAAGCGAATAATACAGACGTGAATAACGCACTGGAATCTTACAAGAAGTTCCTGCTGAAGGAACAGATGTCGCCACT	3600
AS	V V S E S G S D T L E Q S Q P K K P A S T H V G A E S N T I	1227
gp190 ⁿ	T A A G A T A A A A A A A A A T C A	
gp190 ^s	GTGGTGTCTGAATCTGGCTCCGACACACTGGAGCAGTCTCAACCTAAGAAGCCTGCATCTACTCATGTCGGAGCCGAGTCCAATACAATT	3690
AS	T T S Q N V D D E V D D V I I V P I F G E S E E D Y D D L G	1257
gp190 ⁿ	A A A T T A A A A A A A A T A T C A A T T T A A	
gp190 ^s	ACCACATCTCAGAACGTCGACGATGAGGTGATGACGTCTCATTTGTGCTATCTTCGGCGAGAGCGAGGAGTACGATGACCTCGGC	3780
AS	Q V V T G E A V T P S V I D N I L S K I E N E Y E V L Y L K	1287
gp190 ⁿ	A A A A A A A A A A A A A T T T A T T A T T A T A	
gp190 ^s	CAGGTGGTCAACCGGTGAGGCTGTCACCTCTCGGTGATTGATAACATCTGTGCCAAATCGAGAACGAATACGAAGTCTCTATCTGAAA	3870
AS	P L A G V Y R S L K K Q L E N N V M T F N V N V K D I L N S	1317
gp190 ⁿ	T A T T A A G T A A A T A A T A T T T T T A T T A T T C A	
gp190 ^s	CCTCTGGCAGGCTCTATAGGTCTCTCAAGAAACAGCTGGAGAATAACGTGATGACCTTCAATGTCAACGTGAAGGACATCTCTGAACAGC	3960
AS	R F N K R E N F K N V L E S D L I P Y K D L T S S N Y V V K	1347
gp190 ⁿ	A A C T A T A A T C A T A A T T A A A A G T T A	
gp190 ^s	CGCTTTAATAAGAGAGAAAATTTCAAGAACGCTTGGAGAGCGAGTTGATTCCTATAAAGACCTGACCTCCTCTAATACGTTGTCAAG	4050
AS	D P Y K F L N K E K R D K F L S S Y N Y I K D S I D T D I N	1377
gp190 ⁿ	T A T T A	
gp190 ^s	GACCCATACAGTTCTCAATAAAGAGAGAGGATAAATTTCTGTCTAGTTACAATATATCAAGGACTCCATCGACACCGATATCAAT	4140
AS	F A N D V L G Y Y K I L S E K Y K S D L D S I K K Y I N D K	1407
gp190 ⁿ	T A T T A T A A T A T C T A A T T A T A A A A A A A A A	
gp190 ^s	TTCCCTAATGATGTGCTGGGGTATTACAAGATCCTGAGCGAAAAATACAAGTCTGACCTTGACTCTATTAATAAAGTATATCAACGATAAG	4230
AS	Q G E N E K Y L P F L N N I E T L Y K T V N D K I D L F V I	1437
gp190 ⁿ	T A G C T T T A C T T G T A T A T T T T T T A T	
gp190 ^s	CAAGGCGAGAGTGAATAATATCTGCCCTTCTGAATAACATCGAAACCTGTACAAGACAGTGAACGACAAATCGACCTCTTCGTAATT	4320
AS	H L E A K V L N Y T Y E K S N V E V K I K E L N Y L K T I Q	1467
gp190 ⁿ	T T A A A A T A T A T A T A T C A C A A A A A T T A T	
gp190 ^s	CACCTGGAGGCCAAGGTCTCAACTATACCTTACGAGAAGAGCAATGTGGAAGTTAAATCAAGGAGCTGAACCTACCTCAAAACAATCCAA	4410
AS	D K L A D F K K N N N F V G I A D L S T D Y N H N N L L T K	1497
gp190 ⁿ	A T T A T T A	
gp190 ^s	GACAAGCTGGCAGATTTCAGAAAAATAACAATTTCTGCGGAATGCGAGACCTGTCTACCGATTATAACCAACAATCTCCTGACCAAG	4500
AS	F L S T G M V F E N L A K T V L S N L L D G N L Q G M L N I	1527
gp190 ⁿ	C T A G T A T T T T T T C T T A T C T A T T A T A T A T	
gp190 ^s	TTTCTGTCCACTGGCATGGTGTTCGAAAACCTCGCCAAAACAGTCTGAGCAATCTGCTCGACGGCAACCTGCAGGGCATGCTGAACATC	4590
AS	S O H Q C V K K Q C P Q N S G C F R H L D E R E E C K C L L	1557
gp190 ⁿ	A A	
gp190 ^s	TCCCAGCACCATGCGTGAAGAAACAGTGCCTCCAGAAATAGCGGCTGTTTCAGGCATCTGGACGAGCGCGAAGAGTGCAAGTGTCTCCTG	4680
AS	N Y K Q E G D K C V E N P N P T C N E N N G G C D A D A K C	1587
gp190 ⁿ	T T A T T A T T T T C T A T A T A C T	
gp190 ^s	AACTACAAACAAGAAGGAGATAAGTGCCTGGAGAACCCAAACCTACCTGCAATGAAAACAATGGCGGGTGTGACGCCGATGCTAAATGC	4770
AS	T E E D S G S N G K K I T C E C T K P D S Y P L F D G I F C	1617
gp190 ⁿ	A T T C A T A A T T T T T T T T T T T T T T T T T T	
gp190 ^s	ACCGAGGAAGACAGCGGCTCTAACGGAAGAAATACATGCGAGTGTACTAAGCCGAGCTCTATCCACTCTTCGACGGGATTTTTCG	4860
AS	S S S N F L G I F F L L I L M L I L Y S F I * *	1639
gp190 ⁿ	AGTTC C T A A A C A T A T A A T A T A T T	
gp190 ^s	TCCAGCTCTAATTTCTGGGCATCTCTCTGCTGATCCTCATGCTGATCCTGTACAGCTTCATCTAATAGATCGATGG	4940

stop codon Cla 1

THIS PAGE BLANK (USPTO)

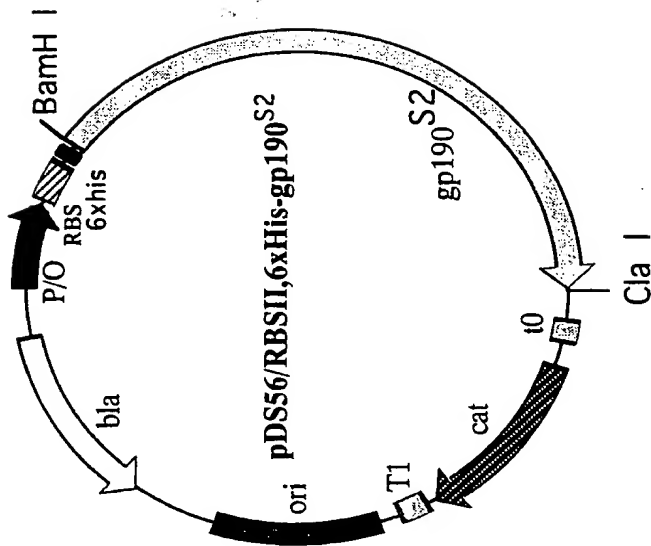
Abb.3D

8 / 16

	N'-terminus	C'-terminus
gp190s1 Sequence		
DNA Sequence	GC <u>ACGCGTATGAAAATC</u> ----- AGCTCTAATTAAATAGGCGGCCGCATCGATGGC	
AA Sequence	Mlu I Met Lys Ile ----- Ser Ser Asn stop codon Not I Cla I	
AA Position	1 2 3 ----- 1619 1620 1621	
gp190s2 Sequence		
DNA Sequence	GC <u>GGATCCGTGACCCAC</u> ----- AGCTCTAATTAAATAGGCGGCCGCATCGATGGC	
AA Sequence	BamHI Val Thr His ----- Ser Ser Asn stop codon Not I Cla I	
AA Position	20 21 22 ----- 1619 1620 1621	

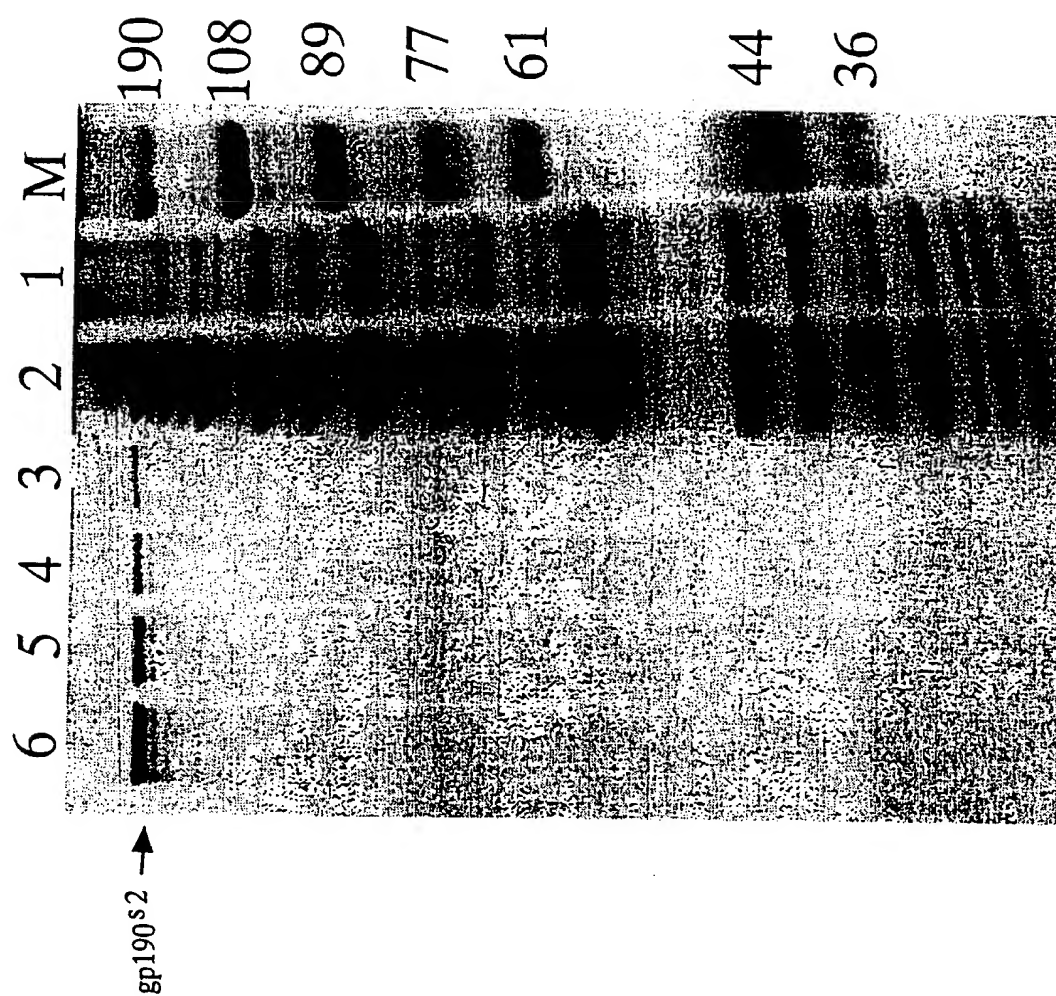
THIS PAGE BLANK (USPTO)

Abb. 4A



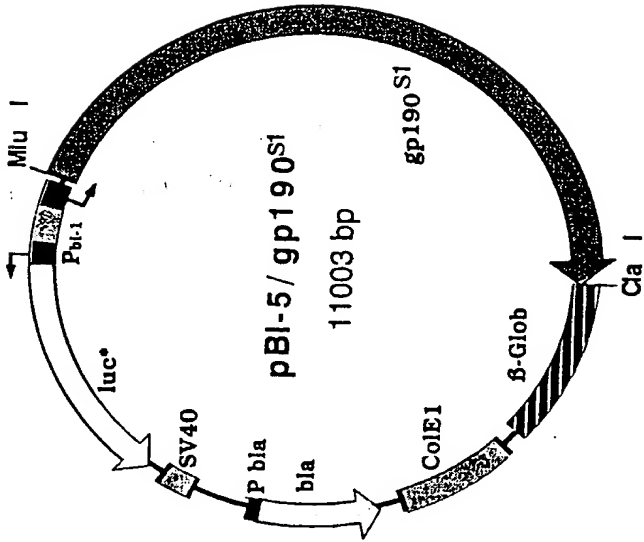
THIS PAGE BLANK (USPTO)

Abb.4B



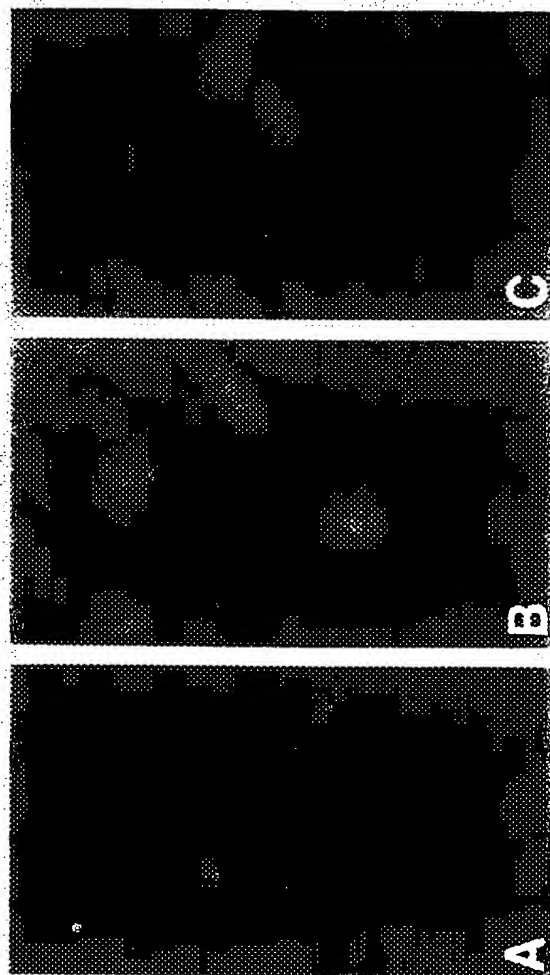
THIS PAGE BLANK (USPTO)

Abb.5A



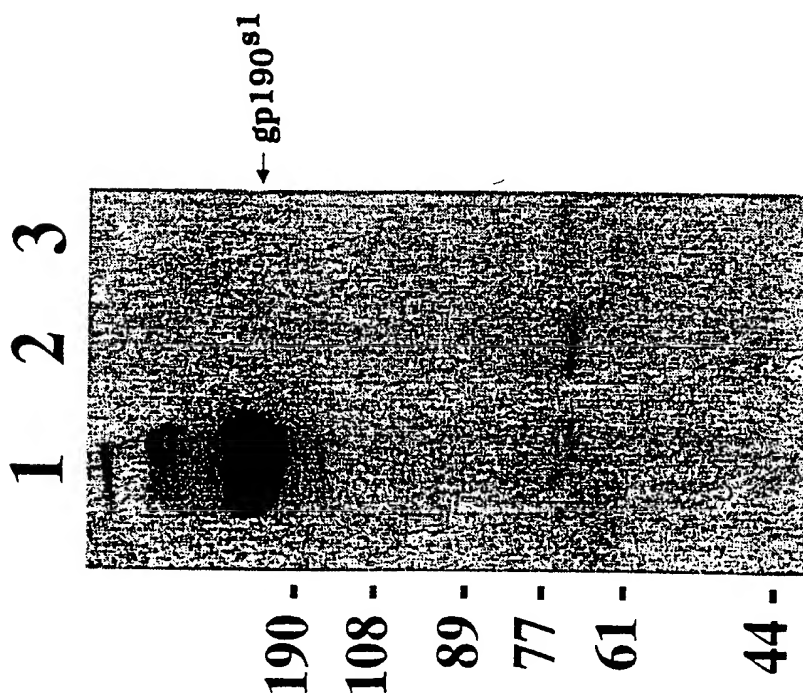
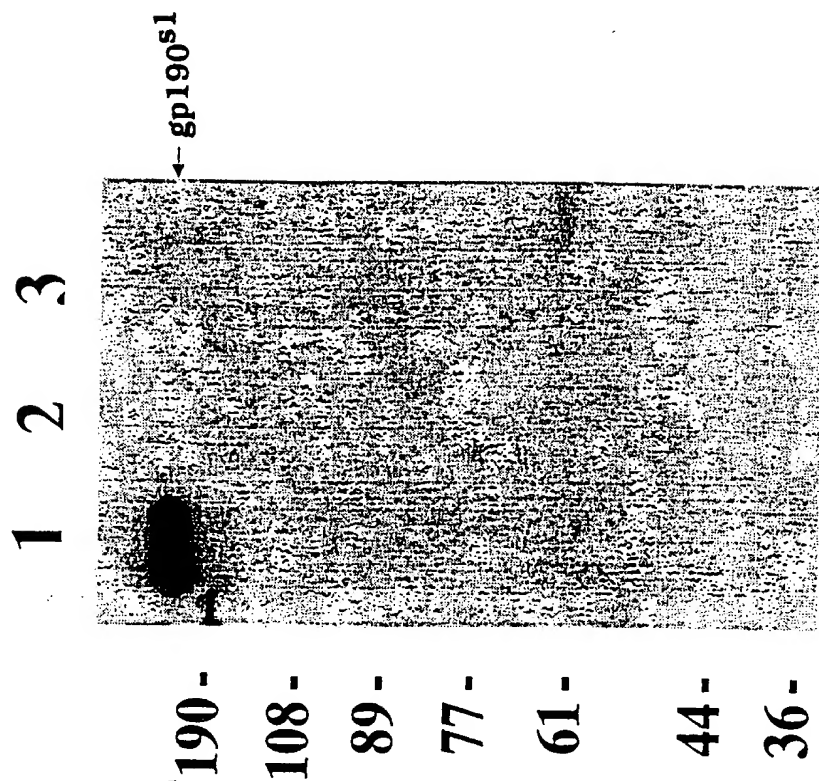
THIS PAGE BLANK (USPTO)

Fig. 5B



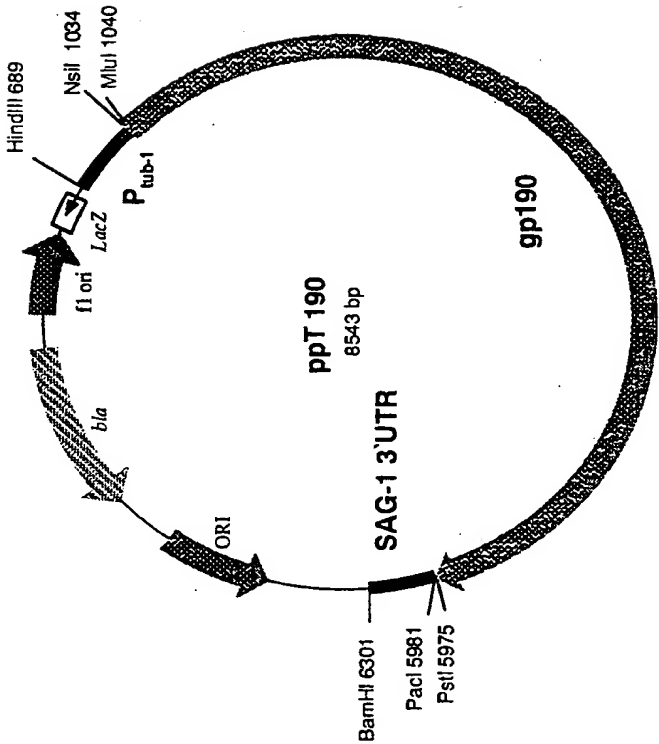
THIS PAGE BLANK (USPTO)

Abb.5C



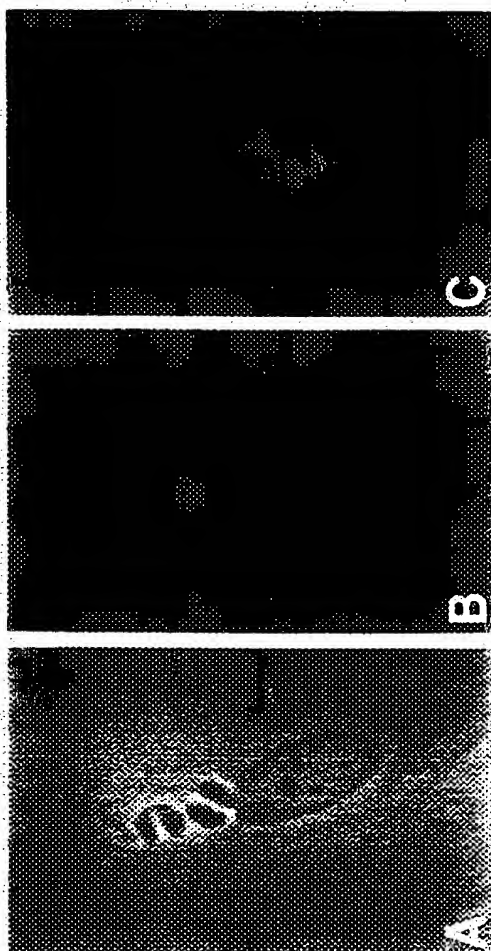
THIS PAGE BLANK (USPTO)

Abb. 6A



THIS PAGE BLANK (USPTO)

Fig. 6B



THIS PAGE BLANK (USPTO)

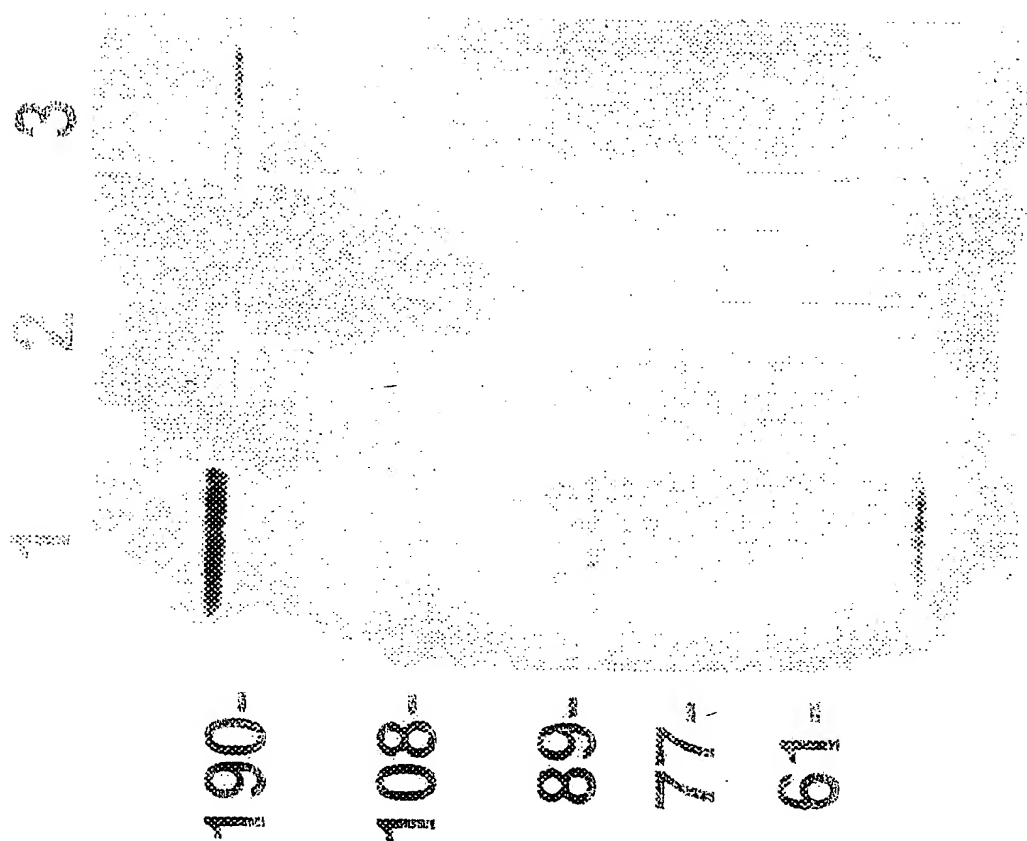


Fig. 6c

THIS PAGE BLANK (USPTO)
